2007 STAR Administration

Scope of Work

March 2, 2006

Table of Contents

Table of Tables

1. Component Task 1: Comprehensive Plan and Schedule for Project Deliverables and Activities (CST, CAPA, CMA, STS, NRT)

1. A. Narrative Schedule

The Standardized Testing and Reporting (STAR) Schedule of Project Deliverables and Activities (see Timeline 1) is included as an attachment to this Scope of Work. This schedule depicts all STAR deliverables and activities for the entire contract period and includes task initiation and completion dates and may be adjusted periodically as needed to reflect new timelines and deliverables.

1. B. Progress Reports

Educational Testing Service (ETS) will communicate all accomplishments to demonstrate the California Department of Education's (CDE) expenditures on the STAR Program by means of monthly and quarterly reports.

Monthly Reports

The monthly report documents ongoing and completed tasks and activities based on the scope of work. In addition, as issues arise during the month, ETS will detail these issues, their causes, and resolutions as part of the monthly and quarterly reports.

The monthly progress report is presented as a detailed narrative invoice from ETS to the CDE. It is sorted by scope of work area and contains a summary section which is the actual invoice, and a section of accomplishments and deliverables that breaks down the costs in the summary section and associates the costs with a specific requirement in the scope of work. The summary includes a history of invoices previously submitted to date.

ETS will submit this report to the CDE by the 15th of the following month. A hardcopy original will be delivered to the CDE. In the event that this report will be delayed beyond the 15th of the following month, ETS will notify the CDE of the expected date of delivery by the 7th of that month.

Quarterly Reports

The quarterly report summarizes Program activity for three months and is accompanied by a CD-ROM that contains portable document format (PDF) deliverables — or the representation of deliverables — for the previous three months. The CD-ROM's files are stored within a directory tree organized by the scope of work section and navigated using an index PDF file, an approach that simplifies the CDE's access to the documentation. ETS will deliver a quarterly report and CD-ROM to CDE staff, and a copy of the quarterly report to State Board of Education (SBE) staff. Both the SBE testing liaisons and the CDE STAR project director will have to approve the content of the quarterly report.

1. C. Management Meetings

ETS will hold weekly management meetings with the CDE and with the STAR team that includes ETS managers of:

- Test Development
- Statistical Analysis
- Computer Technology

- Operations
- Item Banking
- STAR Technical Assistance Center (STAR TAC)
- Appropriate Subcontractor Coordinators

Although each of these managers may hold separate weekly meetings, the purpose of the management meeting is to update and assure that the CDE is informed of all decisions.

For all meetings, including face-to-face and video- or audio-conferenced, ETS will facilitate the meeting, record minutes of the meeting, and track completion of assignments. The minutes will be distributed to the CDE and the entire team within 24 hours of the management meetings.

Annual Meeting. ETS will host an annual three-day meeting in Sacramento which gathers key STAR Team members to meet with CDE program managers and SBE staff members. Those who cannot attend in person may attend via videoconference. The purpose of the meeting is to plan the upcoming year, including detailing any changes to the scope of work and timeline. ETS will provide a draft timeline in MS-Project for all to review. Also, ETS leads a discussion and documents lessons learned to prevent problems from reoccurring and to continue improving the STAR Program both for the SBE/CDE and for all California school districts. The outcome of this annual meeting will be an update to the draft timeline and any changes to the Scope of Work requested by SBE testing liaisons and the CDE STAR project director.

1. D. SBE Meetings and Technical Meetings

Every time the SBE conducts public meetings, ETS project managers and relevant ETS officers are expected to attend to the extent necessary.

1. E. Records and Minutes

At all meetings, including, but not limited to Assessment Review Panel (ARP) meetings (see Section 6), standard-setting meetings, management meetings and item-writing sessions, ETS will take minutes, record information and document any assignments or tasks for follow up. These notes will be formatted in a MS-Excel spreadsheet, as required by the CDE. ETS will keep secure electronic copies of all the records throughout the life of the STAR Program. These minutes are delivered as part of the Quarterly Report deliverables CD-ROM. All minutes will be made available to the SBE upon request.

Each set of minutes will include listings of all those present and their contact information. ETS will review the contact information of attendees to determine if it has changed and update the CDE, if appropriate.

These records will be distributed to the CDE for approval in less than 10 days following each meeting. When approved, all relevant STAR team members will receive copies.

1. F. Report Style

For all reports submitted to the CDE as deliverables, ETS will follow the guidelines established by the *Publication Manual of the American Psychological Association: Fifth Edition* as well as the CDE Style Manual guidelines.

1. G. Quality Control Audit

ETS will perform an annual audit of its work under the STAR Program and provide the results of the audit to the CDE, SBE staff and testing liaisons. This audit will be conducted by the ETS Office of Corporate Quality Assurance (OCQA). This audit will review STAR Program compliance with ETS Standards for Quality and Fairness 2002. Each rigorous standard is

applied to the program-supplied information and documentation in a uniform, systematic, and well-documented process. This annual audit is conducted by ETS staff not affiliated with the STAR Program.

- The audit evaluates requirements of five to eight standards under each of the following chapter headings:
  - Customer Service
  - Fairness
  - Uses and Protection of Information
  - Validity
  - Assessment Development
  - Reliability
  - Cut Scores, Scaling, and Equating
  - Assessment Administration
  - Reporting Assessment Results
  - Assessment Use
  - Test Takers' Rights and Responsibilities

Results of this internal audit will be delivered with the Quarterly Report following the conclusion of the audit cycle.

1. H. Long-term Assessment Plan

ETS will assist the SBE and the CDE in developing a long-term assessment plan. In 2002, the CDE and the SBE published a long-range assessment plan that has been mainly achieved over the last three years.

The six principles articulated in the 2002 long-term assessment plan are still relevant. Within the framework of these principles, ETS will work with the SBE and the CDE to safeguard the validity, reliability and usefulness of the STAR assessments. ETS will work with CDE and SBE staff and testing liaisons to develop key issues and a policy long-range plan to guide policy for the STAR Program.

1. I. Overlap of Contracts and Transition

ETS will do everything possible to assure that transition to the next contract is as smooth as possible.

End of Contract

ETS will deliver all required materials (such as reports, post-test workshops, data files, etc.) to the CDE by December 31 of the year following the last test administration as contracted. One member of the current STAR Program Management Team will serve as a transition manager to assist the new contractor until the end of the calendar year in which the last administration is completed. This transition manager will provide the following deliverables to the CDE or directly to the new contractor, if preferred by the CDE.

The ETS transition manager will have access to key members of the STAR staff from each functional area of expertise to assure that all deliverables listed above and any other reasonable requests are made available to the CDE and the new contractor on a mutually-agreed schedule.

The transition manager will also establish regular meetings with the new contractor during the overlap of the contracts to communicate all that the new contractor may need to know about

the STAR Program. There is no cost to the STAR contract for the provision of these transition activities.

1. J. ETS Management

A senior decision-making person will be officed in Sacramento to serve as a designated liaison to the SBE for the ETS STAR contract. This person will act as a liaison between ETS and the SBE by understanding all policy issues around the STAR Program and its background, and making sure both that the SBE has the information and background they need for policy oversight, and that the SBE's direction and governance are manifest in the program activities and outcomes. This person will have duties to include:

- Learn the policy and legislative underpinnings of the STAR Program and the history and rationale of policy and operational decisions.
- Meet frequently in person with the SBE testing liaisons, the SBE Executive Director, and the other SBE or CDE key staff to determine their information needs and receive guidance as appropriate.
- Provide information and briefings to the SBE liaisons and SBE Executive Director on measurement and evaluation policy issues as needed.
- Keep informed of national trends and developments in statewide assessment programs and act as an additional resource to the SBE to keep them informed of these trends and developments so as to make informed policy decisions.
- Maintain excellent relationships with the CDE leadership so as to foster good communication among ETS, the CDE, and SBE regarding SBE concerns and information regarding the STAR contract.
- Collaborate with ETS Program Management to assure that SBE concerns are reflected in the STAR program direction.
- Keep ETS senior management informed of SBE policy issues and direction and assure appropriate alignment with SBE policies and directives.
- Attend all significant meetings on program operations affecting policy and policy meetings, including all regularly scheduled SBE meetings, and scheduled SBE testing liaison meetings, and other meetings as request by the SBE Executive Director, SBE President, and/or SBE testing liaisons.

This staff position will not be funded through the STAR contract.

2. Component Task 2: Program Support Services (CST, CAPA, CMA, STS, NRT)

2. A. Help Desk

The STAR Technical Assistance Center (TAC) is in place to provide personalized service to district STAR coordinators on STAR Program-related issues.

STAR TAC incorporates all of the required features, including:

- Toll-free telephone and facsimile access
- E-mail access
- Dedicated staff available to respond 7 a.m. to 5 p.m. Monday through Friday (local California time)
- Sufficient staff to assure that all requests for assistance are handled within 24 hours of when they are received
- Maintenance of a log of customer concerns for reporting and categorization purposes
- Providing the CDE with an electronic version of the log within 10 days of a request

ETS will investigate relocating STAR TAC operation to Sacramento in order to provide even higher service levels at a lower cost.

Calls are routed to STAR TAC through Symposium, a contact-center product from Nortel. Symposium allows ETS to track several key statistics to measure performance. This data is available to the CDE for review at any time by request.

ETS will use a short telephone survey to assess operator performance as well as customer satisfaction. This survey is confidential and voluntary. The level of satisfaction with the service provided is rated on a scale of 1 to 5 (1 = poor and 5 = excellent).

Each STAR TAC representative is monitored twice per week by the team leader or director. The results of the monitoring will be entered into NICE, a contact-center product used to evaluate and collect call-quality monitor scores. Representatives' performance is assessed on product knowledge and call mechanics including:

- Using appropriate tone
- Giving complete and accurate information
- Understanding customer needs through appropriate questions or clarifying statements

ETS will maintain a log of its interactions with customers. Interactions with districts are logged into a custom-designed MS-Access database, which has been expressly designed for the STAR Program.

STAR TAC representatives will attend district STAR coordinator events, such as the pre- and post-test workshops.

2. B. Startest.org

ETS will maintain the existing startest.org Web site in a manner that conforms to the CDE's design, accessibility, writing and content, and applications standards as specified in the CDE Web toolkit. Startest.org provides a portal for all of the resources a district STAR coordinator needs to administer the STAR Program.

ETS will, through startest.org, provide district STAR coordinators with timely updates of relevant information. In addition, ETS will respond promptly to requests by the CDE for changes to content and links. The STAR Editor on the ETS Program Management Team will maintain the site content, links, and documents.

Specific improvements to the startest.org site for the new contract period include:

- New sections specific to the Standards-based Test in Spanish (STS) and California Modified Assessment (CMA) with the same look and feel as the current site
- Weekly updated Frequently Asked Questions (FAQ)
- Ongoing review of the site's look and feel to conform to CDE Web site standards
- A built-in survey component that allows district STAR coordinators to provide input into the design and content of startest.org

All changes to the Web site will require CDE approval before implementation.

There will be an annual review of the Web site to assure that it conforms to CDE design, accessibility, information architecture, writing and content, and application standards as specified in the CDE Web toolkit, as well as to evolving Web standards.

In addition to CDE review, ETS will host annual focus groups to continually improve the functionality of startest.org. In addition, STAR TAC will relay feedback from district calls pertaining to startest.org.

2. C. Collection and Monitoring of Information

District Coordinator and Superintendent Contact Information

The STAR Management System will use the County-District-School (CDS) Code Master File to populate its database. Districts will receive from ETS a Superintendent's Designation of District STAR Coordinator Form. The District Superintendent will complete and sign the form and return it to ETS by U.S. Mail or fax. The prior year's district STAR coordinator will also receive a copy of this communication in order to assure receipt and action from the Superintendent. ETS will track the receipt of a completed form for the district along with any updates to the data in the STAR Management System. Any changes to the assigned district STAR coordinator made during a testing year will require a new Superintendent's Designation of District STAR Coordinator Form signed by the District Superintendent. STAR TAC will enter the receipt date of these documents into the STAR Management System, triggering district access to the system. District STAR coordinators will not be able to access the STAR Management System until this form and a Security Agreement have been received from the district.

Security Agreements

District and test site coordinators will receive from ETS the STAR Test (including field tests) Security Agreement for District and Test Site Coordinators (the "STAR Security Agreement form") in June or July together with the Superintendent's Designation of District STAR Coordinator Form. The district STAR coordinator will return a signed Security Agreement via mail or fax. STAR TAC will then record the receipt date of this form. Upon receipt of this form and the Superintendent's Designation of District STAR Coordinator Form, the district STAR coordinator will receive a user name and temporary password to access the STAR Management System.

Charter School Designations

ETS will use the CDS code files to update the STAR Management System to be in synch with the Pearson Educational Measurement (PEM) Management System. In turn, PEM will provide information to ETS on the testing status of each charter school including information on whether PEM has received test materials back from charter schools.

All charter schools identified in the CDE's CDS master file will receive a letter and a Charter School STAR Designation Form similar to what regular districts have in July or August. The form will require the charter school to specify whether it will be testing dependently with the district or whether the charter school will be operating independently of the district for the purposes of STAR testing. ETS will maintain a log of all charter schools to whom forms were sent and of all charter schools returning completed forms along with their independent or dependent STAR testing status. ETS will conduct follow-up telephone, fax and e-mail communications in order to obtain completed forms from all charter schools.

Shipping Information

The STAR Management System will store the shipping address along with an operational calendar and operational hours for each district (including charters). Other data related to shipping (for example, whether the facility has an unloading dock) will also be communicated to the shipping company in order for them to have the appropriate equipment available when making deliveries.

The district will also have the option of providing a secondary shipping address that may be used for the delivery of supplemental orders and reports. For all deliveries other than the initial shipment of materials, the district will have the ability to select whether to receive a shipment at the primary or secondary shipping address.

District STAR coordinators will have the ability to update shipping addresses at any time. The new address is then effective immediately for shipping.

Test Material Orders

The STAR Management System will handle orders in one of the following ways:

- District STAR coordinator will enter online screen data.
- File submission — Districts will either submit online or send via CD-ROM files containing order numbers in a predefined and published interface file format.
- District STAR coordinator will use U.S. mail/fax/e-mail to submit a material order form to STAR TAC staff, then enter the order information directly into the STAR Management System. The district STAR coordinator will verify the information when STAR TAC staff return a copy of the data entry form to review the numbers and confirm the accuracy of the data entry.

Any additional materials requested can be entered as a Supplemental Order. Verified orders will be released for packaging at the appropriate time. All supplemental orders will be released and shipped within a 48-hour (or better) timeframe.

2. D. Terminology

ETS will assure the accuracy and correctness of testing materials including practice tests, test booklets, *Directions for Administration* (DFA), and California Alternate Performance Assessment (CAPA) stimulus cards, manuals, guides, memorandums and e-mail messages, as well as online application screens and startest.org pages. ETS editors follow American Psychological Association (APA) guidelines and the CDE Style Guide.

Additionally, ETS will maintain a terminology list as usage evolves and the CDE adds new terms. See the Terminology List in Table 1.

The Terminology List will not only include words and phrases, but will also indicate proper hyphenation, abbreviation and capitalization.

The ETS/STAR Terminology List will use the CDE Style Guide as its primary resource.

Table 1. ETS/STAR Current Terminology List

AIS	average item score
ARP	Assessment Review Panel
AYP	Adequate Yearly Progress
CAPA	California Alternate Performance Assessment
CAT/6 Survey	California Achievement Tests, Sixth Edition Survey
CDE	California Department of Education
CI	confidence interval
CMA	California Modified Assessment
CRL	California Reading List
CSEM	conditional standard error of measurement
CST	California Standards Test
CTT	Classical Test theory
DIF	differential item functioning
EAP	Early Assessment Program
ELA	English-language arts
EM	expectation maximization
ICC	item characteristic curve
IDEA	Individuals with Disabilities Education Act of 1997
IEP	individualized education program
IRT	item response theory
IUP	Item Utilization Plan
MAD	mean absolute difference
MH D-DIF	equated delta scale (Mantel-Haenszel delta DIF)
NAEP	National Assessment of Educational Progress
NCLB	No Child Left Behind Act of 2001
NR	no response
NRT	norm-referenced test
OCQA	Office of Corporate Quality Assurance

R-FEP	Reclassified-Fluent English Proficient
SBE	State Board of Education
SD	standard deviation
SEM	standard error of measurement
SGID	School and Grade Identification Sheet
SMD	standard mean difference
SPAR	Statewide Pupil Assessment Review
STS	Standards-based Test in Spanish
TAG	Technical Advisory Group
TCC	test characteristic curve
WRMSD	weighted root mean square difference
1PPC	1-parameter partial credit
California academic standards (never abbreviated)	
district STAR coordinator (never DSC)	
double-rated: verb, hyphenated	
English-language arts; but, English-Language Arts CST	
hand-grid: verb, hyphenated	
history-social science (hyphen, not slash)	
scale scores (not scale*d* scores)	
second rater: noun	
Web site (capital W)	
* Capitalize course name when it's part of a test name (Algebra II CST)	
* Do not capitalize course name when it's not part of a test name (a class in algebra II)	

2. E. Workshops for Appropriate Personnel for the CSTs, CAPA, CMA, STS, and NRT

ETS will present:

- Pre- and post-test workshops
- Workshops on new Web applications
- Monthly Web casts

ETS will design materials for 11 pre-test workshops and five post-test workshops. The workshops will offer a train-the-trainer model that provide suggested methods, materials, information, and practice so that district STAR coordinators and County Office of Education (COE) trainers can present their own workshops to site coordinators, who in turn will train teachers and other administrators. The objective of all these workshops is that tests be administered correctly and results be interpreted and used properly.

The pre-test workshops offered in January and February of each year will cover the procedures for receiving, distributing, administering, and returning test materials, with an emphasis on their security. The August post-test workshops will give information and practice in the interpretation of results. California Standards Tests, CAPA, STS, CMA, and norm-referenced tests (NRT) overlap in the pre-test procedures and in similar interpretation of results.

In addition, separate CAPA pre-test workshops will employ a train-the-trainer model for the administration of CAPA, with practice using manipulatives, stimulus cards, giving cues and scoring according to the rubrics, either for Level I or for Levels II–V. Consistency in the administration of CAPA helps assure reliability of the assessments. ETS will issue certificates of completion for the training and templates for district coordinator trainers to issue to CAPA examiners.

Workshop locations. The workshops will be held in a variety of locations across the state. ETS works with County Offices Of Education, which provide facilities for the workshops at no cost to the State.

Description of Materials to be used in the Presentation

ETS will have videos made for pre-test and for CAPA administration.

ETS will create the following materials for pre- and post-test workshops and post them on startest.org:

- Pre- and Post-Test Manuals distributed at workshops or shipped to non-attending district STAR coordinators
- Guides for the STAR Management System, posted on startest.org
- MS-PowerPoint slides
- Quizzes and handouts

ETS will produce a pre-test video and will update it as needed. ETS will also produce a Web cast of the pre- and post-test workshops.

Videotransform, Inc., will duplicate tapes, formatted in both VHS and DVD and distribute them to districts. Starting in 2007, **ETS will distribute all the videos as either VHS or DVD**, based on districts' requests, with enough for each site. At first (2006–07 school year), districts will express their preference for DVD or VHS in an e-mail ordering procedure, and in future years, through the STAR Management System.

Starting in 2007, ETS will arrange for closed captioning, even if it is not available until after workshops begin, and will let districts request these special versions of the videos.

All workshop materials will be approved by the CDE prior to their use.

2. G. California Reading Lists for the CSTs

The California Reading Lists (CRL) can be found on the CDE Web site *www.cde.ca.gov/ta/tg/sr/readinglist.asp.*

ETS will make available the California Reading List on the Internet. ETS will update and revise this list with new reading material, grade-level ranges and annotations as requested by the SBE. Such materials will be added to the list upon the approval of the SBE. The updated California Reading List will then be posted on the California Reading List Web site.

ETS will provide an index that correlates the ranges of pupil scores on the California English-Language Arts (ELA) CST to materials that would be suitable for pupils in Grades 2 through 11.

ETS will also provide a mapping of the California Reading List to the California content standards. This will be delivered as a Technical Report to the CDE during 2006. ETS will make any other changes to the CRL that the SBE or future development requires.

2. H. Released Test Questions

ETS assessment specialists will select 25 percent of the core operational items from each administration for each grade and subject area of the CSTs, CAPA, STS, and CMA for release annually. The numbers of items to be released annually for each grade and subject area are shown in the tables below.

Table 2. CSTs Number of Items to be Released

Test	# of Operational Items per Form	# of items to be released following administration
ELA Grades 2 and 3	65	16
ELA Grades 4–11	75	19
Writing Grades 4 and 7	2	2
Math Grades 2–7 and all specific subject tests	65	16
Science — All grades and subject tests	60	15
History-SS Grade 8	75	19
History-SS Grades 10 and 11	60	15

Table 3. CAPA Number of Items to be Released

Test	# of Operational Tasks per Form	# of items to be released following administration
ELA-— Levels I–V	8	2
Math — Levels I–V	8	2
Science — Levels I–V	8	2

Table 4. CMA Number of Items to be Released

Test	# of Operational Items per Form	# of items to be released following administration
ELA Grades 2 and 3	65	16
ELA Grade 4–11	75	19
Math Grades 2–11	65	16
Science Grades 5, 8 and 10	60	15

Table 5. STS Number of Items to be Released

Test	# of Operational Items per Form	# of items to be released following administration
ELA Grades 2 and 3	65	16
ELA Grades 4–11	75	19
Math Grades 2–7 and all specific subject tests	65	16

ETS assessment specialists will select each set of items for release such that each set of released questions is representative of the various content measured by items on the operational tests. In addition to assuring that the released questions measure a variety of content standards, ETS assessment specialists will make certain that each set of released questions represents the realistic range of difficulty for the items on the test, as well as the variety of types of questions used for measuring various skills.

ETS assessment specialists will use the following criteria to select operational items for release following administration.

- Overall quality of the test item
- Accuracy and clarity of wording and content
- Strong match to the identified content standard and construct
- Statistical reliability of the item
- Range of difficulty
- Representation of a variety of standards and overall blueprint distribution
- Representation of the various components of the standards
- For ELA, any permissions issues pertaining to passages being published on the Web

Because many publishers are reluctant to grant permission for general Web release of their product, such as passages used in ELA assessments, ETS assessment specialists for ELA will choose released items that are associated with passages for which they know Web permissions can be acquired or are not necessary. This will assure the most reliable and representative possible release of ELA items following each operational administration.

For each multiple-choice question recommended for release, ETS will include information about the standard that each item is designed to measure, as well as the correct answer response or item key.

For released writing prompts at Grades 4 and 7, ETS will provide sample student responses illustrating each possible score point in the rubric.

All items, associated data, and sample student responses selected for potential release will be submitted for review and approval to the CDE. The SBE and CDE will approve any and all items for public release. Once items are approved for release, ETS will label them as "released" in the item banks, so that they will no longer be considered for inclusion in any California assessment.

Items will be released only in an electronic format in an HTML format on the CDE Web site.

Communication Plan for Released Test Questions

By June 1, 2006, ETS will submit for approval a communication plan to the SBE staff and test liaisons and CDE regarding Released Test Questions. This plan will include:

- A summary of the types of questions cumulatively released to date and summary of projected number and types of items released in the future;
- Draft print materials addressing the needs of teachers, parents, and the general public regarding understanding of released test questions and how they are useful for understanding the STAR testing program
- Outline of training sessions to be provided by ETS

- A schedule of training sessions subsequent to annual release of test questions. Participants in these workshops will be teachers and local school administrators. The training sessions will also be Web cast and available on the STAR public Web site (see section 2.J) for public viewing.
- A draft letter for SBE/CDE signature that can be sent to statewide organizations (e.g., ACSA, CSBA), legislators, and the general public announcing the availability of the released test questions.

In the development of print materials, ETS will survey other states and incorporate best practices into material development. ETS will also conduct focus groups separately with parents and teachers in conjunction with local training sessions. The purpose of the focus groups will be to obtain feedback on the clarity and usability of materials and guide the design and development of the materials for the coming years.

Both STAR TAC staff and COE STAR liaisons will be trained in handling inquiries about the released test questions.

ETS will be prepared to make a presentation to the SBE at one of their regularly scheduled meetings on Released Test Questions and the results of the ongoing communication plan.

2. I. Communicating About Performance Levels

To enhance understanding of the California testing program, ETS will link the performance levels directly to released test questions (RTQs). This linkage will be based on test statistics and address the question "What kind of achievement is represented by each performance level?"

2. I1. Method to Illustrate Meaning of Performance Levels

In order to give educators, parents, policy-makers, and taxpayers an understanding of the level of knowledge and ability a student must demonstrate to be considered Basic, Proficient, or Advanced, ETS will work with SBE staff and testing liaisons, the CDE, and the ARPs to determine the appropriate data and text that should accompany existing and future released test questions.

As a first step, ETS will gather models from any other states that have linked their RTQs to performance data and present this information to the State as potential models for consideration.

ETS will also present details about how data are linked to National Assessment of Educational Progress (NAEP) released items by ETS and the National Center for Education Statistics (NCES), so that SBE staff and testing liaisons, the CDE, and the ARPs can also use NAEP models as a springboard to decision-making. One NAEP model, a NAEP "item map," is attached to the Scope of Work as representative of the kinds of ideas ETS will present during this first step. The item map would locate released questions on the test scale, to show what kind of achievement is expected for Basic, Proficient, and Advanced students.

A second NAEP model, illustrated in the table below, will also be presented for consideration. This table would give stakeholders an understanding of the percentage of students who answered each released item correctly, again illustrating what proficiency and other performance levels mean.

Table 6. NAEP Model 2

	Multiple Choice Released Test Question			
	Percentage Correct			
Overall Percentage Correct 73	**Below Basic**	**At Basic**	**At Proficient**	**At Advanced**

A third NAEP model is illustrated in the next table. This table is a distractor analysis for released items, showing, by percentage, which incorrect choices were attractive to students. This information also can be disaggregated according to the Basic, Proficient, and Advanced performance levels, again giving stakeholders a better understanding of what kind of achievement is represented by proficiency and other performance levels.

Table 7. NAEP Model 3

	Multiple Choice Released Test Question Percent Choosing Each Answer Choice					
	Average Score	**A**	**B**	**C***	**D**	**Omit**
All Students	60	3	14	60	11	-
Below or At Basic	50	4	17	50	12	1
At or Above Proficient	61	3	12	61	10	1
At or Above Advanced	68	1	13	68	7	-

Following initial presentation of these and other models, ETS will make specific recommendations and present draft samples for consideration by the SBE staff and testing liaisons, the CDE, and the ARPs.

Next, ETS will survey other states and incorporate best practices into material development. ETS will also conduct focus groups separately with parents and teachers in conjunction with local training sessions. The purpose of the focus groups will be to obtain feedback on the clarity and usability of materials and guide the design and development of the RTQ materials.

ETS will then work closely with the SBE staff and CDE to develop a model for final sign-off by all state representatives. After sign-off, ETS staff will create the new RTQ files for distribution and posting on the CDE Web site.

2. I2. Plan for Developing and Distributing the Final Product

ETS will work with the CDE and SBE staff to design and implement a plan for publishing and distributing released test questions, based on the information described in section 2.I1 of this scope of work, that will address the needs and satisfy the questions of California's parents, educators, policy-makers, and taxpayers. Whichever method of release the CDE and SBE prefer, released test questions will clearly describe and illustrate for the public the level of achievement on the content standards that the State expects of proficient students.

Communication Plan for Illustrating the Meaning of Performance Levels

By June 1, 2006, ETS will submit for approval a communication plan to the SBE staff and CDE regarding performance levels. This plan will include:

- A summary of the types of questions cumulatively released to date and summary of projected number and types of items released in the future.

- Draft print materials addressing the needs of teachers, parents, and the general public regarding understanding of released test questions and how they are useful for understanding the STAR testing program.
- Outline of training sessions to be provided by ETS.
- A schedule of training sessions subsequent to annual release of test questions. Participants in these workshops will be teachers and local school administrators. The training sessions will also be Web cast and available on the STAR public Web site (see section 2.J) for public viewing.
- A draft letter for SBE/CDE signature that can be sent to statewide organizations (e.g., ACSA, CSBA), legislators, and the general public announcing the availability of the released test questions.

Both STAR TAC staff and COE STAR liaisons will be trained in handling inquiries about the meaning of performance levels.

ETS will be prepared to make a presentation to the SBE at one of their regularly scheduled meetings on the meaning of performance levels and the results of the ongoing communication plan.

2. J. Assist schools and districts analyze grade-level and course results.

ETS will develop materials and strategies to assist districts and schools in the proper usage and interpretation of STAR Program test scores. These will include:

1. A post-test guide that explains the types of scores produced and what they mean. The outline of this guide is provided in the table of contents of the *2005 Post-Test Guide Technical Information* provided in the following table:

Table of Contents

2005 Post-Test Guide Technical Information

2. A set of post-test workshops each August that provide an in-depth explanation of each component of the STAR assessments. An outline of the 2005 Post-Test Workshop is provided in the following table as a template for future workshops:

Outline
Post Test Workshop 2005

Objective:
District Coordinators will interpret and communicate STAR results.

I. Intro	8–8:30
A. Registration	
B. Welcome and introductions to people	

C. Objectives and agenda	
D. What's new?	
II. Psychometrics	8:30–8:50
A. Year-to-year comparison: Do's and don't's	
B. Interpretations for individuals	
C. Cluster scores	
III. AYP/API Interpretations	
III. Aggregate Reports (including Internet)	
A. What's new?	
B. How to interpret?	
C. Communicate	
1. To administrators	
2. To evaluate programs	
IV. Individual and Teacher Reports	
B. What's new?	
Student report: **How can I use these STAR Program results? More about the STAR Program (if no NRT), Clusters: percent correct, compared with percent correct of proficient students statewide.**	
Labels: CAPA also this year.	
C. How to interpret	
Proficient and above or not	
Areas to focus on	
Not: scaled score of student previous year	
D. How to communicate	
To parents	
To teachers	
VI. Writing test: Identify how writing tests are scored.	
A. Overview of writing in large scale assessments	
B. Readers	
C. Prompts	
D. Anchor sets with training papers	
E. Resources with releases	
VII. Summary	

3. A series of onsite visits to the largest districts in the State as well as individual County Offices of Education (COE) to provide additional consultation about score interpretation. A schedule of district and COE visits will be provided by July of each year to CDE and SBE staff. District visits will be targeted at larger districts; COE visits will be targeted at groups of smaller districts brought together at the COE. The COE visits will be part of the strategy discussed below.
4. Web-based resources, including the *Post-Test Guide* and questions/answers from the post-test workshops; these will be features of startest.org.
5. Web casts of workshops that can be viewed on demand by local district personnel. An outline of such a Web cast is provided below :

Outline

Post-Test Workshop Web Cast 2005

Objective:
District Coordinators will interpret and communicate STAR results correctly.

Topic	**Estimated Time**
I. Intro	
A. CDE staff	7 minutes
a. Welcome	
b. Importance of STAR and results	
c. Encouragement to use results and use correctly	
d. CAHSEE workshop in afternoon	
e. Intro to presenters: ETS Staff	
B. Objectives and agenda	15 minutes
f. E-mail questions to be answered at breaks	
g. Quiz	
C. What's new?	7 minutes
h. Cluster score comparisons on student and teacher reports	
i. Grade 5 science clusters reported	
j. California Reading List Number based on CSTs	
k. Ethnicity added to Subgroup Summary	
l. Braille not indicated on individual reports	
II. Statistical Theory	30 minutes
A. Year-to-year comparison: Do's and don't's	
B. Interpretations for individuals	
C. Cluster scores	
III. Aggregate Reports (including Internet)	60 minutes
A. Common data	
1. Number enrolled	
2. Number Tested	
3. Number Valid Tests	
4. Performance Levels: percent proficient	
5. Mean Scale Score	
B. Reports	
1. Student Master List Summary	
a. By grade	
b. List subjects	
c. Writing abbreviations	
2. End-of-Course Master List Summary: Math and Science	
3. Subgroup Summary	
a. Disabilities	
b. Economic status	
c. Gender	
d. EL Fluency	
e. Ethnicity	
4. Group Summary: CAT/6 Survey	
a. If tested at Grades 3 and 7 only	
b. Compare percentiles to previous years	
5. Unmatched Report	
a. Grade 3, books 1 and 2	
b. Grades 4, 7: writing to MC	
6. Internet	

a. Levels: state, county, district, school, group, subgroup
b. Number and percent of tests
c. Mean scale score
d. Percent at each performance level

C. Communicate
1. To administrators
2. Demo with template
a. Resource: SMLS, Subgroup Sum
b. Average percent correct 2004 vs 2005
ii. By grade
iii. By subgroup
c. What increased, what decreased, what increased the most?

IV. AYP/API — 30 minutes
A. Goal of assessments
B. API: definition, based on what, and how used
C. AYP: API: definition, based on what, and how used
D. Participation rates
1. How to determine what's significant
2. Subgroups
3. How box 26 contributed to participation

Break — 20 minutes
E-mail questions
Question and answer

V. Individual and Teacher Reports — 60 minutes
A. Common data
1. Performance levels
2. Cluster comparison to proficient
B. List and Labels
1. Label: for student's permanent records
2. SML: to see all students in one admin; CAPA in last admin
C. Student reports
1. CSTs
a. Scale score as bar graph: target 350
b. Cluster scores
c. CRL
d. Other resources
2. CAPA
a. Scale score as bar graph: target 35
b. Description of performance levels
D. Teacher reports
1. Summary reports by group (teacher/classroom), grade, math EOC
2. Compares clusters to scores of students in state, district, and to proficient student scores.
3. Compares performance levels across subgroups
4. Compares subgroups this year and last year
5. Other resources
E. How to interpret

1. Proficient and above or not 2. Areas to focus on 3. Not: scale score of student previous year F. How to communicate 1. To parents a. Resources: Translation guide; Appendix A b. Emphasize target of proficient c. Cluster: How close to proficient i. Work on one farthest from proficient 2. To teachers a. Resource: SMLS, Teacher Report b. Cluster: Average percent correct compared to average percent correct by students who scored proficient on test.	
Break	20 minutes
E-mail questions Question and answer	
VI. Writing test: Identify how writing tests are scored. F. Overview of writing in large-scale assessments G. Readers H. Prompts I. Anchor sets with training papers J. Resources with releases K. New layout for 2006	60 minutes
VII. Summary A. Quiz answers	7 minutes
Question and Answer	20 minutes

Role of the County Offices of Education

ETS will provide a "County Office of Education STAR liaison" function in each the 11 county regions in the state. The COE staff member who takes on this function will be trained by ETS to provide training on various technical and operational aspects of the STAR program, including, ordering tests, administering tests, interpreting scores and understanding score reports. The specific duties of a COE STAR liaison may vary from region to region depending on the assessed needs of the districts in that region. By May 2006, ETS will define the role of each of the COE STAR liaisons. These functions and deliverables will be reviewed by SBE staff and liaisons and CDE prior to finalizing. ETS will budget for logistics and training costs to support this collaboration with the County Offices of Education.

During the first quarter of the new contract period, ETS will work with the SBE staff and CDE to formalize a plan for specific STAR activities in each region. ETS will develop a separate scope of work for each of the eleven regions specifying dates, training activities to be provided by COE STAR liaison, support to be provided by ETS, and deliverables.

The goal of the COE STAR liaison will be to extend the training to districts provided by STAR to each district in the state that requests assistance, regardless of size or location.

ETS will use the scheduled pretest and posttest workshops as "train the trainer" opportunities for COE STAR liaisons. ETS will provide all training and master copies of materials needed for COE STAR liaisons to provide training on their own. An ETS program manager will be in charge of the COE STAR liaisons and communicate with them regularly via phone, e-mail and onsite to assure the quality of their training and the information they provide to schools and

districts. In addition to their own training, COE STAR liaisons will have access to STAR staff via videoconference and audio-conference to assist in school and district presentations.

In addition to the COE STAR liaison initiative, ETS will work to leverage the COE special education mission, including the Special Education Local Plan Area (SELPA) network to implement the CAPA portion of training and technical assistance to districts.

Development of a Comprehensive Communication Plan

ETS will support the SBE and the CDE in communicating key STAR Program deliverables (including score results), events, and procedures to school districts.

Communications to Teachers, Parents, and the General Public

ETS will support the SBE and the CDE in communicating the important role of the State's testing program to teachers, parents, and the general public. Specific activities will include:

1. ETS will design and host a Web site to present information about the STAR assessment to parents, teachers and the general public. This Web site will provide such information as:
   a. Released test questions with explanatory text
   b. Links to the Content Standards
   c. Explanations of parent and teacher reports
   d. A "Question and Answer" function to provide specific feedback on specific inquiries
2. ETS will work with the SBE and the CDE to disseminate the print materials widely across the State to parents, teachers, and the general public. Moreover, ETS will work with the SBE staff and liaisons and the CDE to devise a strategy for utilizing the Web site and encouraging more people to use it.
3. ETS will conduct grassroots outreach efforts in areas where the SBE staff and liaisons and the CDE deem appropriate. These efforts would include working with individual districts and parent or community groups to help the targeted populations better understand the testing process.
4. In conjunction with the County Offices of Education, ETS will staff and conduct regional forums to better inform parents and the general public on how to better understand the score reports (and to empower them to make better use of them). These forums would be coordinated to coincide with the release of the STAR results.
5. ETS will work in conjunction with the SBE, the CDE, the Superintendent of Public Instruction, the Office of the Secretary of Education and members of the California State Legislature to further enhance communications and raise awareness of the test by releasing public service announcements.
6. All content of the communications, under the communications plan with districts and the public regarding the STAR Program, must first be approved by the CDE and the SBE staff before being disseminated.

Additional costs to complete these activities are contingent upon approval by CDE and SBE of a contract amendment and upon sufficient funds being made available by the Legislature in future fiscal years.

Communicating Public Opinion Research and Policy

ETS will support the SBE and the CDE by providing key educational policy research and public-opinion survey research related to SBE-defined key issues facing California. ETS will disseminate research and survey data as deemed necessary and appropriate by the SBE. Potential research topics shall be determined by SBE staff and liaisons in consultation with the CDE.

3. Component Task 3: Test Security Measures (CST, CAPA, CMA, STS, NRT)

3. B. Security Breach

3. B1. Plan for Working with School Districts Proactively

ETS will perform security audits at selected test sites throughout the State. The purposes of the audits are to make certain no breaches have occurred and to document breaches that are observed.

Examples of security breaches include, but are not limited to:

1. Removing test materials from testing locations
2. Test examiners' sharing test questions and losing secure test materials
3. Missing test booklets before or after a test administration
4. Photocopying secure test materials
5. Failure to follow published test administration procedures.

Audits will be performed before, during and after the administration of the STAR assessments to monitor how closely site coordinators and administrators are following the required procedures as listed in the *STAR District and Test Site Coordinator Manual* and the *DFA*s.

The ETS Office of Test Integrity will conduct workshops to train auditors to properly conduct all required audits of STAR test sites. These workshops will be conducted in Sacramento and Los Angeles, CA. ETS will conduct additional training via one-on-one phone training and/or Web-based training. The auditors will conduct audits according to the principles and procedures published in Responsibilities of Test Site Auditor and will use the standards prescribed in the *STAR District and Test Site Coordinator Manual* and the *Directions for Administration* for the evaluation. School district test coordinators and superintendents will be notified at least one week in advance of an audit visit.

ETS will plan to conduct 100 audits per year:

- 25 pre-test
- 60 during testing
- 15 post-test

In addition, the Customer Care Coordinators will visit the stated number of sites in these time periods to supplement the audits:

- March (for writing tests): 10 sites
- April: 5 sites
- May: 5 sites
- June: 3 sites
- July: 2 sites

Pre-Test Audits. Audits before test administration will occur one to ten days before the site begins testing. Auditors will examine the locked condition of testing materials and whether examiners have received training. They will also visit classrooms to assure that no teachers have tests (except CAPA administrators who are allowed to review the CAPA Examiner's

Manual during the week before testing). Teachers should not be preparing students for answering STAR questions other than the use of practice tests in Grades 2 through 4.

The visits that occur before the first day of testing will assess test administration planning and test booklet security. The auditor will examine:

1. The handling of test materials at the test site after they are received from the district office
2. Secured storage area where confidential test material is stored
3. Test booklets

Audits during testing. The purposes of auditing during test administration include assessing test administration planning, staff performance, test booklet security procedures, and the testing room environment.

During administration, auditors will visit classrooms to determine whether teachers are:

- Setting up classrooms correctly — all test takers side-to-side facing the same direction with a minimum of three feet spacing, with no materials (posters, chalk content) that may hint at answers.
- Following directions in DFAs and reading "SAY boxes" completely.
- Assuring that the pre-identified or previously-used answer document goes to each correct student.
- Not reading questions, nor answering them for students (except Grade 2 Math questions).
- Collecting all testing materials after each testing period.
- Never leaving testing materials in unlocked rooms, but turning in test materials to the site coordinator each day.
- Assuring that students are not copying or cheating in any way.

Also during testing, auditors will see whether the test site coordinator:

- Accounts for all test materials at all times.
- Locks any test materials not in use.
- Packs scorable materials separately from non-scorable materials.
- Makes arrangements to return all materials to the district no more than two days after testing.

Post-test. Post-testing audits are designed to assure that proper procedures are followed for the return of test materials. Auditors will visit one to two days after the make-up testing period in a school. After testing, auditors will assure that no materials are remaining in the school.

The auditor will:

- Verify that all test booklets were returned to a secure storage location.
- Evaluate the timeliness and adherence to published procedures for packing materials for return shipment.
- Count the test booklets to determine if all were returned.
- Evaluate the process of transporting tests and other testing materials from the test site.

- Evaluate the secure storage of materials at the district before they are picked up by the courier for shipment to the STAR Processing Center.

3. B2. Process for Conducting Investigations of Security Breaches

ETS will conduct an investigation of any security breach that may compromise the STAR Program. The primary task will be to lead any investigation of a confirmed security breach.

An investigator from the ETS Office of Testing Integrity will be available within 48 hours to handle security concerns related to the administration of STAR.

Investigations will include interviews with examiners, students (at the discretion of the district), test site coordinators, and any others who had access to the test booklets. These investigations will attempt to determine the identity of those involved in the incident, recover the missing material, and assess the extent to which the test content was compromised.

If anyone attempts to steal test materials, ETS will:

- Confirm the incident with the examination proctor and others identified in the report
- Interview the test site coordinator
- Submit findings to the CDE

The investigation of security breach reports will include:

- Time and date of investigation
- People interviewed
- Findings of interviews
- Steps taken

By request, the ETS Office of Testing Integrity will conduct an immediate onsite investigation in response to security breaches. As required, CDE approval will be obtained prior to the investigation. Within five days of being informed of a security breach, the ETS Office of Testing Integrity will investigate and report results to STAR management. When necessary, immediate reports will be provided by telephone and/or e-mail.

Reports. Auditors will immediately report any breaches to STAR management who will immediately notify the CDE. Auditors are required to file an online site visit form within three days of the site visit. Customer Care Coordinators will also follow that deadline. The ETS Office of Testing Integrity staff that conduct site audits will be required to file a site-visit form within two days of returning to their office.

The ETS Office of Testing Integrity will review each audit report and summarize the findings to arrive at an overall assessment of the test site or district office. The assessment will be calculated by adding up the points recorded on the evaluation report. The final assessment will be reported as “Acceptable,” or “Improvement Needed.” The ETS Office of Testing Integrity will send the completed summary report to STAR management no later than 10 working days after the test administration.

ETS STAR management will then deliver a monthly executive summary for the CDE and SBE staff. The executive summary will show:

- The number of sites visited in the time period
- A list of the sites and their rating
- An indication of which sites had possible breaches
- A summary and outcome of the breaches

5. Component Task 5: Electronic Item Bank, Data Management, and Documentation (CST, CAPA, CMA, STS)

The timeline for item banking activities can be found in lines 205–242 in the STAR Schedule Project Deliverables and Activities in Section 1. A.

5. A Item Bank

ETS currently provides and maintains the electronic item banks for several of the California assessments including the California High School Exit Examination (CAHSEE), California Standards Tests (CSTs), California Alternate Performance Assessment (CAPA) and Standards-based Tests in Spanish (STS). CAHSEE, CST and STS are currently consolidated in the California Item Banking system. CAPA exists as a stand-alone item bank and will be consolidated into the California Item Banking system. The New Alternate Assessment (California Modified Assessment, CMA) items would be added as a new assessment to the bank.

ETS will work with the CDE to obtain the data for those assessments under contract with other vendors for inclusion into the item bank, using the tools previously developed.

The consolidated item bank will house all CST, CAHSEE, CAPA, CMA, STS and CELDT items and associated statistics by assessment. While ETS will retain ownership of its proprietary software, the item bank and the customized version of the software will be owned and copyrighted by the CDE. The enhanced item banking software will support the full functionality described below.

ETS will redesign the look and feel of all the screens of the item bank in order to make it more consistent with the larger STAR suite of applications. Retrieval of data and movement between screens may also be enhanced, with CDE approval, to improve navigation.

ETS will provide the item banking application using the LAN architecture and the relational database management system, SQL 2000, already deployed. In addition, an MS-Access version of the database will be provided to the CDE.

The software will be consistent with the technical environment of the CDE and interface with all standard MS-Office tools such as Access, Excel, and Word, and will output items to commercial software such as Adobe PageMaker/Illustrator, In-Design, or Corel Draw. The application itself will use standard commercial software tools.

ETS will provide updated versions of the item bank to the CDE on an ongoing basis. ETS will work with the CDE to determine the optimum process if a change in databases is desired.

5. A1. Item Requirements

Data maintained in the electronic item bank will include the following:

- **Unique item identifiers.** Each item will have a unique identifier that is established when the item is first written and is consistent with the item identification system currently in use by the CDE and ETS. This identifier stays the same for all drafts of the item. In addition, inherited legacy items will retain their previous item code in a separate field within the database. Only the current draft of an item exists at any point in time in the item bank.
- **Graphics.** Graphics will be stored in standard formats such as .jpg, with math equations created through MathType. If requested, they will also be delivered in .gif or .tif file formats.

Graphics will be maintained in their native format and will be re-scalable. Graphics files will be stored separately and accessed through the item bank application via file pointers.

- **Items Data, Reading Passages and Copyright.** Specification data associated with each item such as grade, assessment level, and content/expectation classifications will be included in the bank. All item text and graphics, including item stem (with braille versions where appropriate), distractors, and links to artwork and passages; item-format data such as response type (multiple-choice and constructed-response, for example), answer key, and scoring rubric/sample responses; links to item writers and item reviewers; and item review comments or results will be included. Copyright information including permissions, expiration date, and type of usage will be included for all passages. All copyright permissions and expirations will be reported to the CDE for review.
- **Item Statistics.** As field-test and operational forms are administered, ETS will load the statistics into the item bank. The item bank will provide for the maintenance of data as requested and approved by the CDE for all administrations of each item. Examples of data elements included are:

  a. *test form data* such as form designation, position on form, administration date, field-test/operational test designation, and anchor/linking designation

  b. *classical statistics for multiple-choice items*, such as n counts, *p*-value, biserial, point-biserial, biserial and/or point-biserial by distractor, distribution of responses by distractor, omit, not-reached, and double-grid for multiple-choice items

  c. *classical statistics for constructed-response items*, such as mean, standard deviation, and score distributions

  d. *IRT statistics*, such as difficulty, discrimination, guessing, standard error, and model fit

  e. *DIF statistics*, such as focal and reference counts, DIF values, and flags for each focal group

  These will be loaded into the item bank, quality-controlled through summary statistics reporting and manual inspection of sample records, and delivered to the CDE according to the schedule for each assessment.
- **Item status.** The item bank will have both an availability indicator and an item status with an associated date and comment area. The availability indicator identifies whether the item is available for use or not and the status further describes the current state of the item in the process. The status of the item or passage will track its current state in the process, from development through review, field testing, operational use, and release. Rejected items, released items, and items that do not meet statistical specifications will be clearly identified so they are not inadvertently included on subsequent test forms. Statuses and availability are updated programmatically as items are presented for review, reviewed and accepted or rejected, placed on a form for field testing, presented for statistical review, and used operationally. Rejection and release of items requires approval and is monitored and quality controlled.

5. A2. Capabilities

The item bank will provide the capabilities that a test developer or psychometrician might need to maintain a testing program:

- **Item Card.** The item banking software will provide a variety of standard reports in addition to the export of data that is already in place in the software today. An item card report will be produced that will model the standard item card that is printed for assessment review panel meetings. The basic information printed will be the unique item identifier, stem and

distractors. In addition, other components can be selected to print with the item such as a standard set of item statistics, a listing of the associated graphics, manipulatives for CAPA items, and associated passages. A separate passage report will also be available. The exact layout and components will be decided upon in conjunction with the CDE. This feature will be constructed so that additional reports can be added when desired.

- **Item Comments.** Users of the software will be able to add and view comments on individual items. Item comments are available in the current item bank and are populated by ETS with Assessment Review Panel comments for items that are presented for review. In addition, the CDE can add comments to an item. This is the only editable area in the application where data is stored in the item bank.
- **Item Search/Selection.** The item banking software will allow the CDE to search for and select items by all specified criteria including content, test administration, statistical characteristics, and Boolean combinations of data comparisons on all fields in the item bank. Once a group of items is selected, the user will be able to view item data in various ways. Selected items will be displayed on screens approved by the CDE.
- **Form Development, Analysis, and Transfer.** Since items are selected while creating a form, the item bank system will provide running tallies of items and their match to the blueprint, statistical summaries (for example, distributions of *p*-values, point-biserials, keys, cognitive levels and fit), and IRT information, standard error, and test characteristic curves. The curves can be analyzed based on the total number of items in the test or any portion thereof. The CDE will be able to create new forms that conform to appropriate test specifications/blueprints using the form planners which can be exported to MS-Excel by one user, transferred, and imported back into the bank or the replicated bank for use by another.
- **Assessment-specific Requirements.** The California Item Banking system will be able to handle differences between assessments while housing them in the same database application. The software will be able to house classifications and blueprints in different formats, report statistics and statistical fields between assessments in different ways, and display data available in one program but not another, such as manipulatives. The underlying database structure is common to all, although all assessments may not use all data fields. This will allow for differences between items developed for different testing programs including but not limited to standards, frameworks, reported statistics, manipulatives, and item classifications. The application and how the data is used and displayed will be completely configurable. Because of the differences in classification and the uniqueness of each individual assessment, items will not be able to be viewed across assessments. Rather, they are unique and available only within one assessment. The combination of multiple assessments, unique items within an assessment, and the ability to access all assessments individually comprises the consolidated item bank.
- **Item Bank Content Summary.** The software will provide summary analyses and printed reports meeting the CDE's specific needs including identifying the number of items in the bank by classification data such as test, domain, content area, grade/level, reporting category, and standard.

5. A3. Item Bank Delivery

In addition to CST items, the consolidated Item Bank will contain items from the CAPA Program, the CELDT Program and the new CMA and STS components of the STAR Program. CAPA is currently a stand-alone application but will be added as a separate program into the existing item bank, and STS is a new assessment added recently to the consolidated bank as well. CELDT is also a stand-alone application that was populated via an external ETS-

developed process and can be continued in that fashion. CMA will be a new assessment added to the consolidated bank. Updates to the database, application and images will be delivered to the CDE or their authorized agent via CD-ROM.

5. A4. Technical Standards

Data will first be loaded into MS-Excel spreadsheets by the test contractor. Quality controls procedures will then be run against the data to assure accuracy and check for validity before being loaded into the item bank. Then, once loaded, the data is again quality-checked using the same routines that run against the SQL 2000 database. As necessary, ETS will work with the CDE and ETS's test contractors to discuss any changes in format or delivery.

5. A5. Stand-alone Version

A fully functional version of the item bank in MS-Access, configured for a stand-alone PC, will be delivered with every update of the consolidated item bank. There will be no application differences between the multi-user LAN version and the MS-Access stand-alone version. While the application requires a secure login and password, two additional levels of security will be implemented. The first level will be at the CD-ROM level, where the CD-ROM itself is password-protected when it is written. The second level will be at the database level, and the database will be password-protected and encrypted to prevent unauthorized persons from accessing the database even if the user front-end is bypassed. ETS will secure full rights of distribution for run-time components of the tools used in the item bank application outside of MS-Office and Adobe Acrobat.

5. A6. Item Bank Update and Delivery Schedule

Monthly deliveries of data will be scheduled according to test development, administration, statistical analysis and form development schedules for each of the assessments. Updates to assessments will be combined across programs as much as possible. ETS will be responsible for both refining the application during its initial deployment and maintaining the software and the database over the life of the contract. In addition to database updates, subsequent versions of the item bank software that incorporate needed changes identified by the CDE will be delivered. A release methodology will be employed to document and track all updates. This will allow data updates and software changes to be implemented as a single release.

An item bank log will be created that tracks all software changes related to the item bank in a single document by assessment. This document will also contain the release number of when the software change was implemented. In addition, a release document will be prepared that will designate the data updates occurring in each release, along with the software changes that will be delivered. This document will be published on the CD-ROM with each update of the software.

ETS will provide routines and database updates so that only database changes and additions are included in updates. Additional quality control procedures will also be developed to assure that the databases are the same after the replication of changes at ETS and the CDE.

5. A7. Quality Control

ETS will employ extensive quality control procedures, involving both automated reasonability checks and counts and careful manual inspection, to make certain all data are accurate and complete. Quality control processes will also be run against all images to assure that all items have an image in the bank and that all art required by the item is present in the bank. When graphics are modified, all items containing that modified graphic will be reassembled and updated in the item bank.

In addition to testing for specific changes to the database and software before delivery of updates, ETS will execute a regression test to assure that there are no unforeseen problems

as a result of changes made. This same test plan will be created in conjunction with the CDE and ETS test development staff. This test plan will be updated to include software changes as they occur. ETS will provide the results of the testing to the CDE for review.

5. A8. Prepare Summary Reports

ETS will provide the CDE with summary analyses and reports based on data in the item bank. Any data in the item bank is available for reporting and ETS Item Banking staff will work with the CDE to identify the data requested, report format, and timeline. A report request will be designed for CDE use to outline the basic request and timeline. ETS will then have a document from which to begin preliminary design and use as a basis for discussion so that the specific needs are documented and understood in order to deliver the desired outcome.

5. A9. Year-End Data Delivery

The year-end data delivery will be handled in the same fashion as a normal database update and would be a scheduled release and part of the delivery schedule described in Section 5. A6. At this time, all items and statistics would be updated in the item bank. A stand-alone item bank in MS-Access would also be delivered. All documentation including data loading procedures, data delivery specifications, item bank table relationship and field definitions, and user directions will be included in addition to the normal database delivery. Other data formats for export will provided on request.

5. A10. Product Licenses

ETS will secure a perpetual run-time license for all the software that is used in the California Item Banking system and will provide that with the application so that any user with MS-Office can run the application. Those third party licenses will be provided to the CDE during the term of the contract so that the CDE can run the application with MS-Office. The CDE owns the customized version of the software that was designed specifically for the CDE with State funding. In order to modify the code or table structures within the system's currently licensed software, the CDE must purchase developer licenses. These licenses are not transferable and cannot be purchased by any party other than the system owner (CDE). The CDE owns the customized version of the software as developed; however, ETS has a patent pending for the underlying proprietary software that has been previously developed and is owned and copyrighted by ETS. Should this contract transition to a vendor other than ETS, the CDE can continue to use the customized system on a perpetual basis but because of the underlying patent, the competitive vendor cannot resell or reproduce the system.

5. A11. Initial Installation

ETS will complete installation of the fully consolidated item bank by the May 1, 2006 required delivery date. The following are the major tasks that will be carried out to build the initial installation of the item bank and the software application:

- **Item bank specifications.** ETS staff will meet with CDE staff to determine detailed requirements for the enhancements to the item bank and the process for creating, maintaining and updating it, specifically from other contractors. The purpose of these meetings is to determine what if any new data elements should be included in the item bank, what additional software functions are required to maintain and retrieve information from the item bank, design of the item card and other reports, and approval of the new screen style which will follow the other CDE STAR applications. The deliverables from this task will be data element definitions, functional application specifications, reporting specifications, a description of the database maintenance process, and technical documents including a detailed quality plan, deployment plan, communications plan, risk management plan, final project plan and schedule, technology environment architecture, and test plan. These documents will be completed by January 30, 2006 and submitted to

the CDE for approval before software development begins. Approval of these planning and design documents is assumed by February 28, 2006.

- **Application development.** ETS will modify the database (logical and physical design) as agreed upon, design and develop the software modifications needed to maintain the database and retrieve information from it, and thoroughly test the application. The deliverable from this task will be a beta version of the re-designed application to be submitted to the CDE for approval. The design will be based on ETS's proprietary item banking technology and the current customized item banking system owned by the CDE. The beta version will be submitted to the CDE by April 30, 2006.
- **Conduct training and acceptance testing with CDE staff and revise item bank as needed.** ETS will develop an updated user manual for use of the software and will conduct a training session to make certain CDE staff are comfortable using the modified item bank. The deliverable from this task will be a training session, a user manual, and Version 1.0 of the item bank, incorporating any changes needed by the CDE. The scheduling of this task will be determined by the CDE based on their review of the beta version and initial item bank.
- **Versioning and item bank maintenance.** ETS will be responsible for refining the application during its initial use and maintaining the software and the database over the life of the contract. The deliverable from this task will be subsequent versions of the item bank incorporating needed changes identified by the CDE, as well as ongoing maintenance of the data described in Section 5. A.6.

5. A12. Scalable Application

The application is scalable to include as many authorized users as necessary with a maximum of 40 concurrent users. Users will be allowed to have only one instance of the application open at a given time, but they will be able to move between assessments based on the access granted by their User ID.

5. A13. Compatibility with CDE Environment

Before implementing any changes to hardware, software or network environments, ETS will obtain approval from the CDE to assure that the changes are compatible with the CDE environment.

5. A14. Item Bank Security

The measures ETS takes for assuring the security of electronic files are as follows:

- Electronic forms of test content, documentation, and item banks are backed up electronically, with the backups kept offsite to prevent loss from system breakdown or a natural disaster.
- The offsite backup files are kept in secure storage, with access limited to authorized personnel only.
- To prevent unauthorized electronic access to the item bank, advanced network security measures are used.

The electronic item banking application will include a login/password system to authorize access to the database or designated portions of the database. In addition, only users authorized to access the specific database will be able to use the item bank. Users will be authorized by a designated administrator at the CDE and at ETS. SBE staff or liaisons will be provided access upon request.

5. B. Documentation

Existing data dictionaries will be reviewed and updated with any enhancement to data noted in the documentation. Data dictionaries contain information about every field in every table in both the item bank and in the flat files used to populate the database from other contractors. The data dictionaries reside in MS-Word or Excel and include the table, field name, data definition, valid values and description of every data element. Data will adhere to the CDE's Data Resource Guide unless other approval is received from the CDE. This document will be a deliverable from the item bank specifications meeting with the CDE and will be provided at least one month prior the delivery of any application or database.

When change requests are approved and modifications are made to the software, this document will be updated and delivered with the updated application.

ETS will develop a user manual and will conduct training sessions for CDE staff as needed to make certain they are comfortable using the item bank. Revised user manuals and additional training will be provided as the item banking system is updated and revised on a schedule to be determined by the CDE. Technical support will also be available by telephone as needed.

6. Component Task 6: Item and Task Development (CST, CAPA, CMA, STS)

All items developed by ETS will meet the technical criteria established in the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) Standards for Educational and Psychological Testing.

The ETS Commitment to Quality Item Development

ETS assessment directors will assure that all APA standards are met, that items and tests meet ETS quality standards, and that items align to the California content standards.

ETS item and passage writers will be thoroughly trained in the California content standards and in ETS's item-writing and passage-writing principles.

At least two ETS assessment specialists will carefully review and edit each item for technical quality (for example, one right answer, clearly stated stem, absence of clueing, plausibility of distractors), match to standard, and conformity with California-approved item-writing practices.

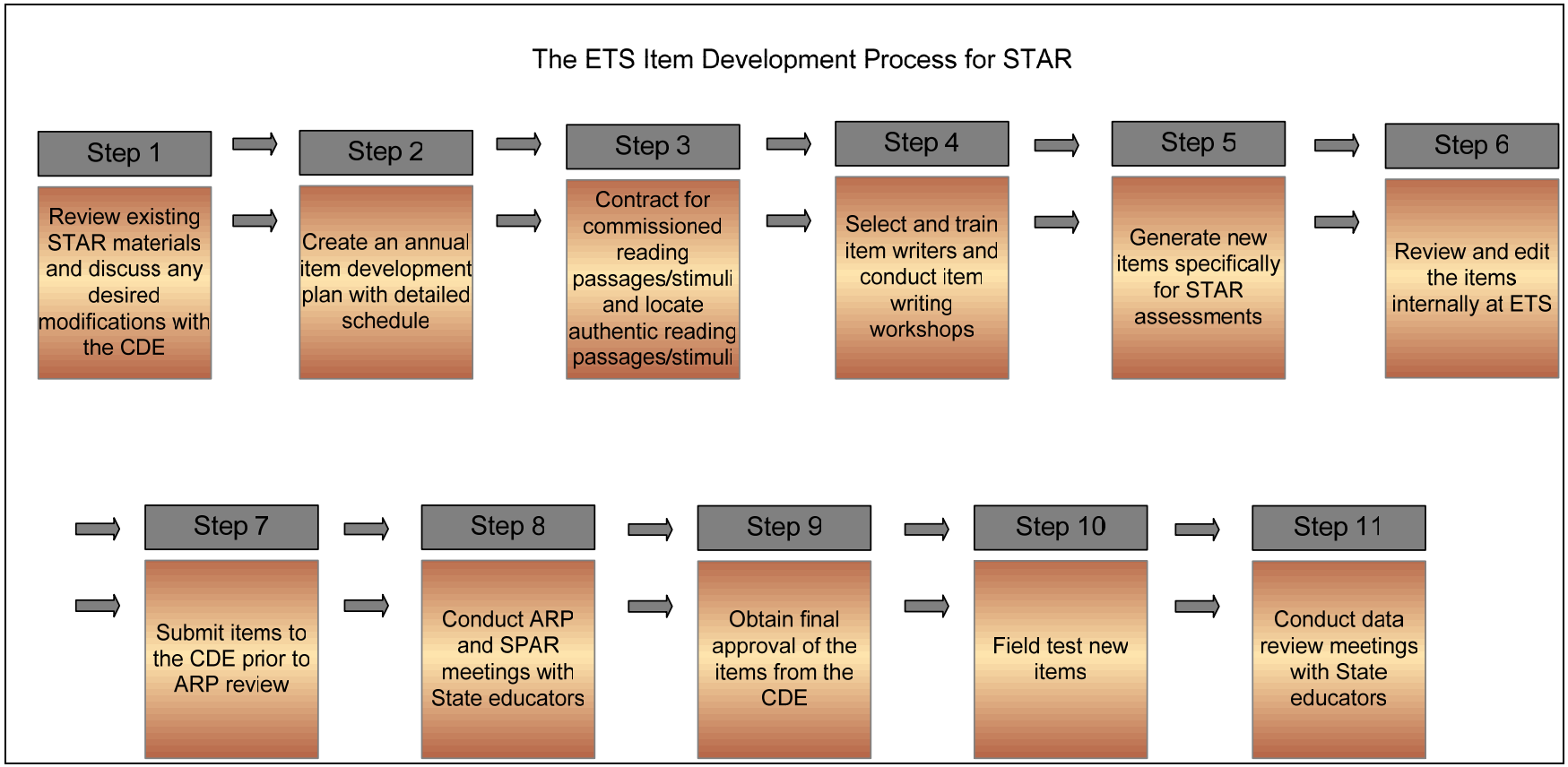
- ETS's trained fairness reviewers will evaluate each passage and item for bias and sensitivity issues.
- ETS senior content staff will also review every item based on their experience with K–12 assessments and their understanding of the California standards.
- ETS editorial staff will evaluate the items for clarity of expression, suitability of language for the grade level, and adherence to style guidelines.
- ETS copyeditors and proofreaders will check each item for grammatical and typographical correctness.

Only after this series of internal reviews will items be submitted to the CDE and ARPs for their review. ETS will discuss with the CDE and SBE staff and testing liaisons any modifications to the current processes that they believe will further increase quality and efficiency.

6. A. Overall CST, CAPA, and STS Item Development

In conjunction with the SBE staff and liaisons, the CDE and California educators, ETS will develop items for the STAR assessments (including Grade 8 and Grade 10 *NCLB* science tests) by annually undertaking the development steps shown in the following figure. The timeline for the item development activities is included in lines 247–273 of Project Deliverables and Activities, which can be found in the attachment labeled "Timelines".

Figure 1. Item Development Process



Each of these steps is discussed briefly below and detailed in the following sections.

1. Review existing STAR materials and discuss any desired modifications with the SBE staff and liaisons and the CDE. The STAR Test Development team will maintain the Item Specifications documents so that they are consistent with the most recent versions of the test blueprints and reflect the wishes of the ARPs, the SBE, and the CDE. ETS will continue to revise and refine the item specifications documents throughout the contract period.
2. Create an annual item development plan with detailed schedule. To create the development plans for each assessment, ETS assessment specialists will meet with the CDE assessment staff and determine development needs for the program. ETS assessment specialists will take detailed notes and share these notes with appropriate ETS staff, as well as other CDE counterparts. ETS will seek approval from SBE staff and liaison and the CDE for its plans for developing items. The ETS item development plans will outline in detail, with specific start and completion dates, the schedule for the year's development, including the CDE review schedule (with sufficient time for CDE staff to evaluate all items), and the timetable of committee meetings. The development plan will be reviewed and revised during the summer and fall of each contract year. It will form the foundation for all ETS development work and will be given to ETS staff as well as the CDE, who will know what must be accomplished each week for the successful completion of the development plan.
3. Contract for commissioned reading passages/stimuli and locate authentic passages/ stimuli. Passages for the Reading and Writing sections of the ELA tests must be engaging, well written, and directed at the appropriate grade level. Reading passages will be given rigorous ETS internal review to attempt to make certain that all materials meet the appropriate STAR Program requirements before they are submitted to the CDE. ETS evaluates passages against the following criteria:
    - Overall Quality
    - Appropriateness of Content
    - Diversity
    - Variety
    - Reading Level
4. Select and train item writers. Conduct item-writing workshops. ETS assessment specialists will oversee the item-writing process and assemble and train highly qualified teams of writers to create the STAR items for each assessment.
5. Generate items specifically for STAR assessments. All items during all phases of the development process will be aligned to the California content standards, conform to approved test blueprints and test item specifications, and follow the STAR Style Manual for the CDE.
6. Review and edit the items internally at ETS. After trained ETS or California item writers write the items, there will be a series of comprehensive internal reviews to evaluate and verify the overall quality of the test items before they are prepared for presentation to the CDE and the California committees.
7. Submit items to the CDE prior to ARP review. After the items have received internal ETS approval, they will be submitted to the CDE prior to the ARP meetings. ETS will submit a precise schedule showing the number of items to be delivered, the content areas, and the scheduled dates.

8. Conduct ARP and Statewide Pupil Assessment Review (SPAR) meetings with State educators. ETS assessment specialists will facilitate the item review sessions.
9. Obtain final approval of the items from the CDE. Following the completion of the review meetings, ETS will seek final approval from the CDE on all items. ETS recognizes that the CDE retains the final decisions about all items.
10. Field test new items. Field testing items is one of the most critical steps in the development process. It is essential that an adequate number of high-quality items result from field tests to support the ongoing validity of the STAR assessments.
11. Conduct data review meetings with State educators. Once items have been field tested, ETS will prepare the items and statistics for review by the ARPs. ETS assessment specialists will facilitate the data review sessions with qualified psychometric staff on hand for technical assistance. Upon completion of the meeting, ETS will provide the SBE staff and the CDE with summaries of the recommendations based on the field-test analyses and committee reviews that are relevant to future form construction and item banking for the STAR assessments. All final decisions on acceptance of items will rest with the CDE in consultation with SBE staff.

6. B. Test Specifications

The test specifications for each current CST will be submitted to CDE staff for review and approval.

The test specifications perform the following functions:

- Define the content of the test (i.e., the specific standards being tested)
- Indicate the test blueprint (the numbers of items for each standard)
- Show the psychometric properties of each item from its most recent administration (whether field-test or operational)
- Provide the properties of the overall test form, including the mean b-value and mean point-biserial for the test

In addition, the test specifications reflect the desired psychometric properties for each item (for example, the upper and lower limits for the b-value).

The test specifications are also summarized in the annual Technical Report.

ETS will revise the test specifications for the CSTs and CAPA at the CDE's request on an annual basis prior to test construction or whenever there is a change in a test blueprint. When changes do occur, ETS will make appropriate technical adjustments to assure comparability over time for successive cohorts of students.

For the STS and the CMA, ETS will develop new test specifications for review by the CDE and the SBE staff, prior to approval by CDE. The test specifications for the STS will be equivalent in rigor to the CSTs. For these new tests, ETS will submit proposed test specifications and item parameters prior to construction of any operational forms and form planners, allowing at least 15 days for CDE review.

The test specifications include content descriptions that show what dimensions of knowledge, skills, processes, and standards are assessed by the test. As described above, the test specifications show the standards being tested, the numbers of items per standard, and the statistical parameters of the items and the test. A second ETS document, called the Item Specifications or the Instructions to Item Writers, gives additional information about what knowledge and skills the items are designed to assess.

Because item specifications for the CSTs and CAPA already exist, ETS's primary responsibility will be to maintain the specifications in order to remain consistent with the direction of the CDE and the SBE.

Currently, the item specifications include, for each standard tested:

- A full statement of the standard;
- The number of items on the blueprint for the standard;
- The reporting cluster in which each standard is found;
- The components of each standard that can and should be tested by the totality of items for that standard in the item bank; this description includes the dimensions of knowledge, skills, and processes that are assessed by the items written for the standard.
- A description of the kinds of item stems appropriate in multiple-choice items for the standard and sample stems for the standard;
- A description of specific kinds of items to be avoided, if any (for example, no items about insignificant details in a passage);
- For ELA, specific guidelines for the different genres of reading passages to be used for the standard; these guidelines include a list of topics to be avoided, acceptable ranges for length, expected distribution of passages by genre, guidelines for readability and concept load, and expected use of artwork.

For STS and CMA, the ETS Test Development Project Lead, along with the assessment specialists, will work with the CDE and the State's ARPs to draft specifications for these new tests using the criteria above.

For STS, the same item specifications will be used as for the CSTs, with sample stems added in Spanish. In this way, the degree of rigor called for in STS items will be the same as the degree of rigor called for in CST items.

Item specifications are not meant to limit item writers' creativity. Rather, these documents describe the general characteristics of the items for each content standard, indicate item types or content to be avoided, and define the content limits for the items, thereby promoting efficiency and consistency in item development.

ETS will submit all recommended edits to existing specifications, and will submit drafts of new specifications to the CDE for review and approval. After the edits have been incorporated, the CDE will have 15 working days to approve the edits before the documents will be submitted to the CDE for final approval.

6. C. Item Utilization Plan

ETS has developed an Item Utilization Plan to continue the development of items for CST, STS, CAPA, and CMA over the next five years. This plan includes strategies for continued coverage of all appropriate standards for all tests in each content area and at each grade level. The tables below show the numbers of items that are required for development of operational forms for the CST, CAPA, STS and CMA annually.

Table 8. Number of Operational Items Required Annually for the CST Assessments

CST			
Test	**Core Forms per Year**	**# of Operational Items per Form**	**Subtotal of Operational Items per Content**
ELA Grade 2	1	65	730
ELA Grade 3	1	65	
ELA Grade 4	1	75	
ELA Grade 5	1	75	
ELA Grade 6	1	75	
ELA Grade 7	1	75	
ELA Grade 8	1	75	
ELA Grade 9	1	75	
ELA Grade 10	1	75	
ELA Grade 11	1	75	
Writing Grade 4	2	1	4
Writing Grade 7	2	1	
Math Grade 2	1	65	910
Math Grade 3	1	65	
Math Grade 4	1	65	
Math Grade 5	1	65	
Math Grade 6	1	65	
Math Grade 7	1	65	
General Mathematics	1	65*	
Algebra I	1	65	
Algebra II	1	65	
Geometry	1	65	
Summative High School Mathematics	1	65	
Integrated Mathematics 1	1	65*	
Integrated Mathematics 2	1	65*	
Integrated Mathematics 3	1	65*	
Science Grade 5	1	60	660
Science Grade 8 *NCLB*	1	60	
Science Grade 10 *NCLB*	1	60	
Biology	1	60	
Chemistry	1	60	

CST			
Test	**Core Forms per Year**	**# of Operational Items per Form**	**Subtotal of Operational Items per Content**
Physics	1	60	
Earth Science	1	60	
Integrated Science 1	1	60*	
Integrated Science 2	1	60*	
Integrated Science 3	1	60*	
Integrated Science 4	1	60*	
History-SS Grade 8	1	75	
History-SS Grade 10	1	60	195
History-SS Grade 11	1	60	
TOTALS	**42**	**2497**	
* Items on the General Mathematics and the integrated forms are repeated from other tests.			

Table 9. Number of Operational Items Required Annually for the CAPA Assessments

CAPA			
Test	**Core Forms per year**	**# of Operational Tasks per form**	**Subtotal of Operational Tasks per Content**
Level I ELA	1	8	
Level II ELA	1	8	
Level III ELA	1	8	40
Level IV ELA	1	8	
Level V ELA	1	8	
Level I Math	1	8	
Level II Math	1	8	
Level III Math	1	8	40
Level IV Math	1	8	
Level V Math	1	8	
Level I Science	1	8	
Level III Science	1	8	32
Level IV Science	1	8	
Level V Science	1	8	
TOTALS	**14**	**112**	

Table 10. Number of Operational Items Required Annually for the CMA Assessments

CMA			
Test	**Core Forms per year**	**# of Operational Items per form**	**Subtotal of Operational Items per Content**
ELA Grade 2	1	65	730
ELA Grade 3	1	65	
ELA Grade 4	1	75	
ELA Grade 5	1	75	
ELA Grade 6	1	75	
ELA Grade 7	1	75	
ELA Grade 8	1	75	
ELA Grade 9	1	75	
ELA Grade 10	1	75	
ELA Grade 11	1	75	
Math Grade 2	1	65	650
Math Grade 3	1	65	
Math Grade 4	1	65	
Math Grade 5	1	65	
Math Grade 6	1	65	
Math Grade 7	1	65	
Math Grade 8	1	65	
Math Grade 9	1	65	
Math Grade 10	1	65	
Math Grade 11	1	65	
Science Grade 5	1	60	180
Science Grade 8	1	60	
Science Grade 10	1	60	
TOTALS	**23**	**1560**	

Table 11. Number of Operational Items Required Annually for the STS Assessments

STS			
Test	**Core Forms per year**	**# of Operational Items per Form**	**Subtotal of Operational Items per Content**
ELA Grade 2	1	65	730
ELA Grade 3	1	65	
ELA Grade 4	1	75	
ELA Grade 5	1	75	
ELA Grade 6	1	75	
ELA Grade 7	1	75	
ELA Grade 8	1	75	
ELA Grade 9	1	75	
ELA Grade 10	1	75	
ELA Grade 11	1	75	
Math Grade 2	1	65	910
Math Grade 3	1	65	
Math Grade 4	1	65	
Math Grade 5	1	65	
Math Grade 6	1	65	
Math Grade 7	1	65	
General Mathematics	1	65	
Algebra I	1	65	
Algebra II	1	65	
Geometry	1	65	
Summative High School Mathematics	1	65	
Integrated Mathematics 1	1	65*	
Integrated Mathematics 2	1	65*	
Integrated Mathematics 3	1	65*	
TOTALS	**24**	**1640**	
* Items on the General Mathematics and the integrated forms are repeated from other tests.			

ETS will submit a biannual report to the CDE and the SBE on the status of the item banks and item development. This report will outline the numbers of items in the banks that are available for forms construction and the numbers of items that are being developed and field tested each year for inclusion in the item banks.

6. C1. Release of 25 Percent of Each CST Annually

The Item Utilization Plan assumes that 25 percent of the operational items for each CST assessment will be released each year. ETS recognizes the importance of providing this information to students, parents, and teachers, as well as the need to maintain an item pool of sufficient quantity to build future operational forms. The item development plan described below in Section 6. C3 anticipates the release of 25 percent of the items per year in each subject area at each grade level, leaving enough items in the pool to construct operational forms each year. This release plan will allow the item pool to grow slightly even with the attrition of items that have become dated or have not performed as expected. Section 2. H described the communication plan surrounding the release of test questions.

6. C2. Items Proposed for Replacement Annually

The Item Utilization Plan assumes that five percent of the operational items for each assessment will need to be replaced each year because of normal attrition. Taking into account the 25 percent release of items, the five percent expected attrition, the need to pinpoint development to "critical" standards (standards in which items tend not to perform well statistically year after year), and the customary refreshment rate of 50 percent for each operational form, ETS has planned the following replacement percentages by content area for the CSTs. The numbers of items shown represent items to be field tested. For all content areas except Science, at least 75 percent of all field tested items are expected survive with usable item statistics. For Science, the expected survival rate is 60 percent survival, especially in the early years of the new *NCLB* tests.

Table 12. Replacement Percentages for the CSTs

Content Area	Percentage of Operational Form to be Replaced per Grade or Course	Number of Items to be Field Tested per Grade or Course
English-Language Arts	110%	84 items
Math	65%	42 items
Science	200%	120 items
History-Social Science	56%–70%	42 items

The plan calls for a greater percentage for ELA than math, for example, because most ELA items are based on passages. If a sufficient number of items fails to survive, all items associated with a passage are lost. There will be an even greater replacement percentage for Science because Science has a greater number of critical standards than the other content areas. The percentages for both Math and History-Social Science will provide enough items to cover release, attrition, and specific development to critical standards.

For CAPA, ETS will field test 150 percent of each operational form for each content area each year. Each operational form contains eight tasks with plans to field test 12 tasks per year. This proportion would allow for a 25 percent item release and a five percent attrition rate, while gradually increasing the overall size of the CAPA item bank.

For the new STS grades as well as the CMA, ETS will conduct a census field test of as many items as the size of the testing population can accommodate, allowing for approximately 2,000

responses per item. In subsequent years, ETS proposes a 50 to 100 percent replacement percentage, depending upon the size of the testing population. These percentages will be refined with the CDE in consultation with the SBE staff and liaisons as the size of the populations become better known.

6. C3. New Items to be Developed, Field Tested, and Delivered for each CST, CAPA, and STS

The Item Utilization Plan described below for each test represents ETS's understanding of the numbers of items that need to be developed and field tested each year in order to sustain forms construction over the next five years. While numbers for the CMA have been included, they may change following SBE approval of the test blueprints.

California Standards Tests (CST)

ETS will slightly increase in development and field testing of items for each CST assessment over the 2003 approved item utilization plan on the Web site. This increase will provide two benefits:

1. To provide more development within specific critical standards (for example, the syllabication standard in Grade 2);
2. To enable more precise form construction at the reporting cluster level in each test. Constructing forms with greater comparability at the reporting cluster level allows a greater level of comparability in forms from year to year.

This plan does not include new development or field testing for the General Mathematics, Integrated Mathematics, or Integrated Science assessments, nor does it include field testing for Writing prompts.

The following table gives the numbers of items ETS will develop for CST field testing for each of the next three years. As mentioned earlier, development will be directed toward the twin goals of 1) replacing items lost to release and attrition; and 2) increasing the numbers of items in the critical standards.

Table 13. Number of Items to Field Test Annually for the CST Assessments

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
ELA Grade 2	14	6	84
ELA Grade 3	14	6	84
ELA Grade 4	14	6	84
ELA Grade 5	14	6	84
ELA Grade 6	14	6	84
ELA Grade 7	14	6	84
ELA Grade 8	14	6	84
ELA Grade 9	14	6	84
ELA Grade 10	14	6	84
ELA Grade 11	14	6	84
Math Grade 2	7	6	42
Math Grade 3	7	6	42

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
Math Grade 4	7	6	42
Math Grade 5	7	6	42
Math Grade 6	7	6	42
Math Grade 7	7	6	42
General Math	1	6	NA
Algebra I	7	6	42
Algebra II	7	6	42
Geometry	7	6	42
Summative High School Mathematics	7	6	42
Integrated Mathematics 1	1	6	NA
Integrated Mathematics 2	1	6	NA
Integrated Mathematics 3	1	6	NA
Science Grade 5	20	6	120
Science Grade 8 *NCLB*	20	6	120
Science Grade 10 *NCLB*	20	6	120
Biology	20	6	120
Chemistry	20	6	120
Physics	20	6	120
Earth Science	20	6	120
Integrated Science 1	1	6	NA
Integrated Science 2	1	6	NA
Integrated Science 3	1	6	NA
Integrated Science 4	1	6	NA
History-SS Grade 8	7	6	42
History-SS Grade 10	7	6	42
History-SS Grade 11	7	6	42
TOTALS	**379**	**228**	**2226**

Writing Prompt Development

The fall 2005 writing prompt field test produced 14 prompts at Grade 4 and 16 prompts at Grade 7. At the rate of using two prompts per year per grade, there are currently sufficient prompts available for STAR through the 2009 testing cycle. Those writing prompts that were not selected by the ELA ARP from the fall 2005 field test will not be reworked for further item prompt development.

California Alternate Performance Assessment (CAPA)

Table 14 shows the numbers of tasks to be field tested annually for each level and subject area.

Twenty-five percent of the operational items for CAPA will be publicly released, just as items are released for the other STAR tests.

Table 14. Number of Tasks to Field Test Annually for the CAPA Assessments

Test	FT Versions	# of Field-Test Tasks per Form	Total # of New FT Tasks per Year
Level I ELA	6	2	12
Level II ELA	6	2	12
Level III ELA	6	2	12
Level IV ELA	6	2	12
Level V ELA	6	2	12
Level I Math	6	2	12
Level II Math	6	2	12
Level III Math	6	2	12
Level IV Math	6	2	12
Level V Math	6	2	12
Level I Science	6	2	12
Level III Science	6	2	12
Level IV Science	6	2	12
Level V Science	6	2	12
TOTALS	84	28	168

ETS special education experts will facilitate an item-writing workshop, where California educators familiar with the student population will create the tasks on site. The ETS CAPA development team will review and edit the tasks as appropriate prior to submitting them for CDE and ARP review.

Standards-based Tests in Spanish (STS)

For development of the new STS, ETS will add several grades to the development process each year until all grades are fully operational. Table 15 outlines the plans for initial development, field testing, and operational testing for the STS assessments.

Table 15. Development Plan for STS

Test	2005	2006	2007	2008	2009
RLA Grade 2	Begin Development	Initial field testing	First operational testing	–	–
RLA Grade 3	Begin Development	Initial field testing	First operational testing	–	–

Test	2005	2006	2007	2008	2009
RLA Grade 4	Begin Development	Initial field testing	First operational testing	–	–
RLA Grade 5	–	Begin Development	Initial field testing	First operational testing	–
RLA Grade 6	–	Begin Development	Initial field testing	First operational testing	–
RLA Grade 7	–	Begin Development	Initial field testing	First operational testing	–
RLA Grade 8	–	–	Begin Development	Initial field testing	First operational testing
RLA Grade 9	–	–	Begin Development	Initial field testing	First operational testing
RLA Grade 10	–	–	Begin Development	Initial field testing	First operational testing
RLA Grade 11	–	–	Begin Development	Initial field testing	First operational testing
Math Grade 2	Begin Development	Initial field testing	First operational testing	–	–
Math Grade 3	Begin Development	Initial field testing	First operational testing	–	–
Math Grade 4	Begin Development	Initial field testing	First operational testing	–	–
Math Grade 5	–	Begin Development	Initial field testing	First operational testing	–
Math Grade 6	–	Begin Development	Initial field testing	First operational testing	–
Math Grade 7	–	Begin Development	Initial field testing	First operational testing	–
Algebra I	–	–	Begin Development	Initial field testing	First operational testing
Algebra II	–	–	Begin Development	Initial field testing	First operational testing

Test	2005	2006	2007	2008	2009
Geometry	–	–	Begin Development	Initial field testing	First operational testing
Summative High School Mathematics	–	–	Begin Development	Initial field testing	First operational testing

For initial field testing of each STS grade or course, ETS will conduct a stand-alone census field test for each new grade as it is developed. Table 16 shows the number of items to be field tested for each grade and course of the STS for the initial census field testing. Because Grades 2, 3, and 4 have been field tested under the previous scope of work, they are not included in the census field test. The field-test plan includes a number of linking items on each form. The linking items will allow the items to be equated to a common scale.

Table 16. Plan for Census Field Tests for STS

Test	Stand-alone Census FT Versions	# of Unique Items per Form	# of Linking Items per Form	Total # of New FT Items
RLA Grade 2	NA	NA	NA	NA
RLA Grade 3	NA	NA	NA	NA
RLA Grade 4	NA	NA	NA	NA
RLA Grade 5	8	75	16	600
RLA Grade 6	5	75	16	375
RLA Grade 7	5	75	16	375
RLA Grade 8	5	75	16	375
RLA Grade 9	5	75	16	375
RLA Grade 10	5	75	16	375
RLA Grade 11	5	75	16	375
Math Grade 2	NA	NA	NA	NA
Math Grade 3	NA	NA	NA	NA
Math Grade 4	NA	NA	NA	NA
Math Grade 5	8	65	16	520
Math Grade 6	5	65	16	325
Math Grade 7	5	65	16	325
General Mathematics	1	65*	16	NA
Algebra I	5	65	16	325
Algebra II	5	65	16	325
Geometry	5	65	16	325
Summative High School Mathematics	5	65	16	325

Test	Stand-alone Census FT Versions	# of Unique Items per Form	# of Linking Items per Form	Total # of New FT Items
Integrated Math 1	1	65*	16	NA
Integrated Math 2	1	65*	16	NA
Integrated Math 3	1	65*	16	NA
TOTALS	**80**	**1240**	**288**	**5320**
* Items on the General Mathematics and the integrated forms are repeated from other tests.				

Following the initial census field testing for each grade, STS field testing will continue as embedded field testing, just as occurs for the CSTs. Table 17 shows the numbers of items to be field tested each year for the STS. The variance in the number of versions for both the census field test and the annual embedded field tests reflects the variance in student n-counts at the different grade levels.

Table 17. Number of Items to Field Test Annually for the STS Assessments Subsequent to Census Field Testing

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
RLA Grade 2	12	6	72
RLA Grade 3	12	6	72
RLA Grade 4	12	6	72
RLA Grade 5	8	6	48
RLA Grade 6	5	6	30
RLA Grade 7	5	6	30
RLA Grade 8	5	6	30
RLA Grade 9	5	6	30
RLA Grade 10	5	6	30
RLA Grade 11	5	6	30
Math Grade 2	12	6	72
Math Grade 3	12	6	72
Math Grade 4	12	6	72
Math Grade 5	8	6	48
Math Grade 6	5	6	30
Math Grade 7	5	6	30
General Mathematics	1	6	NA
Algebra I	5	6	30
Algebra II	5	6	30

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
Geometry	5	6	30
Summative High School Mathematics	5	6	30
Integrated Mathematics 1	1	6	NA
Integrated Mathematics 2	1	6	NA
Integrated Mathematics 3	1	6	NA
TOTALS	**152**	**144**	**888**

ETS will develop items for the STS in the same manner in which they are developed for the CST. This includes writing, reviewing, and field testing the correct number of items per standard to make certain that forms can be constructed to the same blueprints and statistical targets for operational testing from year to year. To serve that end, ETS will develop numbers of items per standard for each subject area and grade in proportion to the test blueprints, as shown in the tables below.

New Alternate Assessment (CMA)

For development of the new CMA, ETS will add several grades to the development process each year until all grades are fully operational. Table 18 outlines the plans for initial development, field testing, and operational testing for the CMA assessments.

Table 18. Development Plan for CMA Mathematics, English-Language Arts, and Science

Grade	For 2006	For 2007	For 2008	For 2009
2	Development	Field test	Operational	Operational
3	Development	Field test	Operational	Operational
4	Development	Field test	Operational	Operational
5	Development	Field test	Operational	Operational
6		Development	Field test	Operational
7		Development	Field test	Operational
8		Development	Field test	Operational
9		Development	Field test	Operational
10		Development	Field test	Operational
11		Development	Field test	Operational

Like the STS, the CMA field testing will begin with a stand-alone, census field test. This field test will take place in the spring of 2007 and include newly developed items and newly purposed items from the CST item bank.

Field testing will occur at all grades so that appropriate item statistics can be obtained from the actual test-taking population. Table 19 shows the numbers of items to be field tested for the initial CMA census field test. The field-test plan includes a number of linking items on each form. The linking items will allow the items to be equated to a common scale.

Table 19. Plan for Census Field Tests for CMA

Test	Stand-alone Census FT Versions	# of Unique Items per Form	# of Linking Items per Form	Total # of New FT Items
ELA Grade 2	5	43	22	237
ELA Grade 3	5	43	22	237
ELA Grade 4	5	53	22	287
ELA Grade 5	5	53	22	287
ELA Grade 6	5	53	22	287
ELA Grade 7	5	53	22	287
ELA Grade 8	5	53	22	287
ELA Grade 9	5	53	22	287
ELA Grade 10	5	53	22	287
ELA Grade 11	5	53	22	287
Math Grade 2	5	43	22	237
Math Grade 3	5	43	22	237
Math Grade 4	5	43	22	237
Math Grade 5	5	43	22	237
Math Grade 6	5	43	22	237
Math Grade 7	5	43	22	237
Math Grade 8	5	43	22	237
Math Grade 9	5	43	22	237
Math Grade 10	5	43	22	237
Math Grade 11	5	43	22	237
Science Grade 5	5	38	22	212
Science Grade 8	5	38	22	212
Science Grade 10	5	38	22	212
TOTALS	**115**	**1054**	**506**	**5776**

Following initial census field testing, items for the CMA be field tested in embedded forms similar to the CST and STS assessments. Table 20 shows the numbers of items to be field tested annually for the CMA following its initial development and census field test.

Table 20. Number of Items to Field Test Annually for the CMA Assessments

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
ELA Grade 2	8	6	48
ELA Grade 3	6	6	36
ELA Grade 4	4	6	24

Test	FT Versions	# of Field-Test Items per Form	Total # of New FT Items per Year
ELA Grade 5	4	6	24
ELA Grade 6	4	6	24
ELA Grade 7	4	6	24
ELA Grade 8	4	6	24
ELA Grade 9	4	6	24
ELA Grade 10	4	6	24
ELA Grade 11	4	6	24
Math Grade 2	8	6	48
Math Grade 3	6	6	36
Math Grade 4	4	6	24
Math Grade 5	4	6	24
Math Grade 6	4	6	24
Math Grade 7	4	6	24
Math Grade 8	4	6	24
Math Grade 9	4	6	24
Math Grade 10	4	6	24
Math Grade 11	4	6	24
Science Grade 5	8	6	48
Science Grade 8	6	6	36
Science Grade 10	6	6	36
TOTALS	**112**	**138**	**672**

The development plan outlined above assumes the approval of the blueprints as they currently exist for the CSTs.

For development of the new writing items for the CMA, ETS will develop items using the same development process currently being used to develop writing items for the California Standards Tests. Items will be written to meet the requirements of the item writer specifications for the English-Language Arts CMA. Table 21 outlines the plans for initial development, field testing, and operational testing for the writing test components of the CMA.

Table 21. Key Dates in the Development of CMA Writing Items

Writing items are created	Spring 2007
CDE/ARP review of writing items	Summer 2007
Writing items, group 2, are created	Spring 2008
CDE/ARP review of writing items, group 2	Summer 2008
Field test of writing items	Fall 2008

Scoring field test	Fall 2008
CDE/ARP review of data from field test	Fall 2008
First operational administration of writing items	Spring 2009
Field test of writing items, group 2	Fall 2009

Six writing items will be field-tested in the fall of 2008 and again in the fall of 2009 for each of grades 4 and 7, in order to provide a yield of at least two items for operational use for the following spring operational test administration. Pending approval by the ARP and CDE, ETS proposes that all writing items be developed in the same genre — narrative— at both grades.

Field test and operational books will follow the same format and length as the CST writing books. This means that the field test will be in two books, with a certain number of multiple-choice items included in each form as a set used for equating. The Writing Prompt and Response booklet will be scannable, while the booklet with the multiple-choice questions will not.

Because the estimated pool of test takers is less than 20,000 at each grade, the sample size per field test item will be about 1,750. Until a pool of writing items is in the item bank, a writing field test must occur every year at each of the two grades, due to the limited size of the pool of test-takers. Writing items will be part of the released test question sets that will be prepared for the CMA.

Prior to operational use, items with field test data will be presented to the CDE/ARP for review. Rangefinding will occur prior to the operational use of these items, as it does on the CST writing tests.

Additional costs to complete CMA writing activities are contingent upon approval by CDE and SBE of a contract amendment and upon sufficient funds being made available by the Legislature in future fiscal years.

6. C4. Method for Achieving Comparable Year-to-Year Results

All tests for the STAR Program will be assembled to rigorous content and statistical specifications to obtain tests that are as parallel as possible. Tests will be assembled to IRT target-test characteristic curves and conditional standard-error curves to assure that the tests perform comparably from year to year. The ETS Test Development teams will monitor the item pools annually to identify critical standards for which new item development is most needed. Section 7. A3c details ETS's plans for assuring that forms are comparable from year to year.

6. D. Reporting Cluster Reliability

Several CST reporting clusters have a small number of items in them. This fact necessarily limits the reliability of the cluster scores. To increase the reliability of these reporting clusters, ETS will make more complete use of the item information in the current tests to increase the reliability of individual student cluster scores.

Individual Student Cluster Reports

Given the structure of the CSTs, and the fact that they are scaled with the Rasch item response theory model, ETS will use the Objective Performance Index (OPI) procedure to increase the reliability of cluster scores for individual students. The OPI procedure has been demonstrated to be accurate.

The OPI procedure, which is calculated for each reporting cluster for each student, provides a more accurate estimate of the student's percent correct on that cluster than the student's

observed percent correct on the cluster. The OPI takes advantage of the fact that while each reporting cluster supplies unique information, performance across the clusters is correlated.

The OPI is a weighted average of two pieces of student data:

1. **The student's actual percent correct on a reporting cluster**
2. **A prediction of the percent correct for that cluster based on the student's total test performance**

The standard OPI procedure will produce a cluster score that is expressed in percent correct units, along with a confidence interval for that score. Expected scores for students just reaching the Proficient and Advanced levels can also be displayed on the same report as they are now.

School, District, and State Cluster Reports

Experience has shown that when OPI scores are accumulated for groups that are the size of a classroom or greater, the results are very similar to accumulations of observed percent correct scores for the clusters. Thus, the OPI procedure does not affect the reliability of school, district or State cluster reports.

6. E. Item and Task Writing for the CSTs, CAPA, CMA, and STS

ETS content area assessment specialists, under the direction of its Assessment Directors, will oversee the item-writing process and assemble experienced, highly qualified teams of writers to create the items for the STAR assessments.

6. E1. Item Writers

In addition to utilizing its current pool of California item writers, ETS will conduct a dual-purpose item-writer training workshop to be held during each of the first two years of the new contract for the purpose of recruiting and training additional California educators to write items for the CST and STS assessments. Because it is important that the STS items be developed with the same rigor and to the same specifications as the CST items, ETS will train writers for both tests at the same time using the same materials. ETS will provide content staff to work with writers for both assessments.

For the CMA, ETS will use a subset of the CST writers in the ETS pool —- those writers who also have experience working with special education students. These individuals will write items for Grades 2 and 3 in Reading and Math and for Grades 5, 8, and 10 in Science. While the CMA items are technically the same as CST items in most grades, the special nature of the test, along with a lack of low-level Grade 2 and 3 items in the CST bank, calls for an independent item development effort for the lower grades. CMA items to be field tested for Grades 4 through 11 will be developed from several sources, including items currently in the CST item banks.

CAPA tasks will be written during an item-writing workshop to be held in Sacramento each year. During these workshops, California educators with experience working with special education students will be recruited to work with ETS staff onsite to write new tasks for ongoing CAPA assessments.

Items written for all tests will meet and exceed the highest standards for quality. Many of the writers selected will have experience working on CST and CAPA items and tasks from the current contract with ETS.

ETS will also contract with individuals to locate and write reading passages and write language passages. The passages used in the ELA sections of the tests will be well-written, interesting, age-appropriate, and rich with opportunities for test items.

All item writers and passage locators for the CST assessments will have the minimum of a bachelor's degree or a teaching credential in their content area. In addition, ETS will require that the writers and locators have at least three years of classroom teaching experience at the appropriate grade level.

In addition to the qualifications outlined above, writers for the CAPA and CMA will also need to have at least two years of experience working with special education students.

For the STS, writers will need to meet all of the qualifications for the CST writers, as well as be bilingual and biliterate in Spanish and English.

In selecting the pool of writers, ETS will also take special care to include an appropriate representation of writers representing the diversity of the California population. The CDE will be provided with the names and résumés of all item writers selected to write items for the program. The CDE may choose not to use a particular item writer based on the strength of his or her résumé.

ETS will recruit item writers in the following manner. First, ETS will write an invitation letter with CDE approval that will be sent to every school superintendent. The superintendents will be invited to nominate teachers or other educators to the ARP, and nominees will be asked to respond by filling out an application and attaching a résumé. The applications and résumés are then reviewed by ETS assessment specialists and then by CDE staff members, who make the final selections. Accompanying the invitation letter will be a clear explanation of the process envisioned for item writing and the benefits to the program from implementing this plan.

ETS will include a specific item development exercise as part of the application materials. Educators who express interest by sending in the application and résumé will receive this exercise as a final part of the application process. ETS will use a short document containing three parts: 1) general training about aspects of technical quality; 2) poor items needing revision; and 3) a small item-writing assignment to bring to the first meeting. ETS will develop this screening exercise under the direction of CDE staff.

6. E2. Guidelines

Item writers will receive several documents to help them understand the intricacies of the CST, STS, CAPA, and CMA.

These documents will include:

- An overview of the California content standards for each specific test, subject area, and grade
- A copy of the blueprints for each of the appropriate assessments
- A review of general guidelines for item writing. The general guidelines ETS uses are collected in the Guidelines for Item Writers; ETS has developed its Guidelines for Item Writers for all K–12 participants in the assessment development process
- A copy of the Item Specifications document for the test, subject area, and grade
- Confidentiality and non-disclosure forms
- Information on how to write items to avoid problems of bias and sensitivity
- A checklist for item writers
- Any other materials requested by the CDE

With the use of these materials, as well as frequent interaction and communication with the ETS content development staff, item writers will become familiar with the format and content required for the development of items for each of the California STAR assessments.

The ETS Guidelines for Item Writers includes information on developing well-written items that are designed to be fair and defensible, have content validity, and mirror sound, proven instructional practices. Item writers will be asked to meet the following criteria for all items developed for the STAR assessments:

- Match to content standard
- Match to construct
- Appropriate difficulty level
- Excellent technical quality
- Appropriateness to purpose
- Freedom from problems of bias or sensitivity
- Appropriate language complexity

Copies of the ETS Guidelines for Item Writers will be provided to the CDE.

6. E2a. Content Standard Alignment for the CSTs and STS

ETS assessment specialists have received guidance on the interpretations of the standards from both the CDE and the panel members who participate in item reviews. This information is incorporated into the Item Specifications, enabling item writers to understand how the California Academic Content Standards should be assessed. For the STS, most of the information captured in the CST specifications will apply from the beginning. New information may be added if needed. For CMA, this information will be gathered throughout the development process as it has been for the CSTs and CAPA.

6. E2b. Content Standard Link for the CAPA and CMA

Items and tasks developed specifically for CAPA and CMA will be linked to grade-level standards, though they may not measure them in precisely the same way CST and STS items will.

In the Item Specifications, the constructs for each content area are clearly defined so that item writers and internal reviewers understand the CDE's expectations. The match to standard is just one verification made during the extensive internal reviews conducted by the assessment specialists prior to sending the test items to external reviews. If an item is a poor match to the standard during the internal review, the item is revised or discarded. However, if a poor match to the standard is discovered after field testing, the item will remain the property of the CDE.

6. E4. Attrition

As is outlined in Section 6. C, Item Utilization Plan, the numbers of items and tasks field tested for the CST, STS, CAPA, and the CMA will be sufficient to maintain and perhaps increase the current level of validity and reliability of the assessments from year to year, but not excessive enough to include high rates of attrition and thereby unnecessary costs.

6. E5. Item/Task Writing Meetings

ETS will hold the combined CST and STS item-writing training workshops in January of each year. The item-writing training workshops for CAPA will be held in April of each year. Following the yearly item-writing training workshops, ETS will send updates via hard copy and e-mail to all new and previously contracted California item writers regarding any changes in the

documents used at the workshop. Changes may include modifications of the blueprint or an interpretation of a standard, newly suggested stems, or stems that are no longer acceptable to the CDE. All materials will be sent to the CDE for review at least 30 days prior to the item-writing workshop, and updates for each year will be sent 30 days prior to mailing the materials to item writers.

ETS assessment specialists will begin the item-writing workshops with an introductory presentation and training, using the materials included in the item-writing workshop packet described in Section 6. E2. After training, writers will divide into groups and will then begin writing several items. ETS assessment specialists will work with the writers and provide feedback as items are written. The two-day training workshop will be based on the following agenda, subject to approval by the CDE.

After the item-writer training, ETS will assess the items written during the training and then make item-writing assignments based on the strengths of each writer.

ETS will execute the full range of meeting logistics and will process all committee member payments. Logistical arrangements include:

- Contracting for suitable meeting rooms that are well-lighted, have plenty of working table space, and are comfortable for working groups
- Arranging for catering services for the meetings, including continental breakfast, morning and afternoon break refreshments, and lunch
- Arranging for meeting equipment for the work of the committees
- Submitting names of committee members to the CDE for approval
- Contacting potential review-committee participants to announce meetings and to arrange for travel
- Reimbursing committee participants per California State travel regulations for travel, lodging, and per diem
- Reimbursing districts for substitute costs for teacher participant
- Assuring the security of all materials

ETS will work with the CDE to determine appropriate dates for the meetings.

6. E6. Internal Review

After trained item writers write the items, ETS will conduct a series of comprehensive internal reviews to evaluate and verify the overall quality of the test items before they are prepared for presentation to the CDE and the California review committees.

The ETS process for review includes:

- An internal content review
- An internal editorial review
- An internal bias and sensitivity review
- Senior review, including external reviewers for validation of content on an as-needed basis

Internal Content Review

Every item receives at least two content reviews by the lead ETS STAR content area assessment specialist and one other equally qualified colleague.

ETS uses the following guidelines for these reviews.

- Match of each item to the identified standard
- Accuracy of the content of the item
- Match of each item to the principles of high-quality item development
- Adherence to the Principles of Universal Design
- Difficulty of the item
- Relevance of each item as the item relates to the purpose of the test
- Readability of the item
- Appropriateness of any artwork, graphs, figures, etc.

These reviews make sure that the test items are in compliance with ETS standards, as well as specific California requirements. These reviews also focus on the accuracy of the content of the passages and items. The internal reviews conducted by ETS assessment specialists include content checks involving one or two sources. In addition, ETS science assessment specialists will review reading passages that focus on science, and social studies assessment specialists will review reading passages that focus on historical or cultural concepts.

ETS will implement the Principles of Universal Design in assessment, as published by the Center for Universal Design at North Carolina State University, and create art with consideration to large-print and braille requirements. Except when State standards require testing skills that cannot easily be rendered into braille (for example, interpreting scatter plots or evaluating use of typographical aspects of texts), all the items ETS develops will be translated into braille.

The artwork and graphics for the items will be created during the internal content review period so that the assessment specialists can evaluate the correctness and appropriateness of the art early in the item development process.

After evaluating each item against these criteria, the reviewers will accept the item as written, suggest revisions or recommend that the item be discarded. At this point, item writers receive feedback on their items so they may learn the nuances of the program and better understand the importance of using the Test Specifications and blueprints when formulating questions.

Internal Editorial Review

After the assessment specialists and the manager review each item, specially trained editors will review each item in preparation for review by the CDE and the review committees. These experienced editors check questions for clarity, correctness of language, appropriateness of language for the grade level, adherence to style guidelines, and conformity with acceptable item-writing practices. Editors often query the assessment specialists about clarity of content or suggest greater precision in the wording.

Language clarity is the focus of both content and editorial reviews. Each item ETS develops is clear and precise in both wording and concept. Items will be designed to assess knowledge in only the subject area being tested.

Internal Bias and Sensitivity Review

Only ETS staff members who have participated in ETS Fairness Training, a rigorous, internal ETS requirement, conduct the next level of review. These staff members have been trained to identify and eliminate questions that contain content or wording that could be construed as offensive to or biased against members of specific ethnic, racial, or gender groups. These trained staff members review every item before it is prepared for the CDE and committee review. Items that do not meet the criteria are revised or discarded.

Senior Review

As a final quality-control step, the assessment director or another "senior reviewer" reads each item before it is ready for the external review process. This step will bring years of training and experience and "fresh eyes" to search for any problems that may have been overlooked by staff members. Additionally, on an as-needed basis, ETS uses external reviewers to validate content, particularly in the science and history content areas. This crucial review, both internal and external, is designed to assure a quality product and helps ETS maintain its high percentage of acceptance rates with State departments and State review committees.

6. F. CDE Review

Following the rigorous ETS internal review process, all newly developed items and tasks for the CST, STS, CAPA, and CMA assessments will be sent to the CDE for review and recommended changes. The items will be presented to the CDE one per page, with the text of the standard reprinted at the top along with additional classification data like the item code, the correct answer, and the cognitive level, so that CDE staff can easily evaluate the item.

A preliminary schedule for CDE reviews of 2006 item development is shown in the next table. ETS anticipates that CDE reviews of new items will occur on or around the same time each year for each development cycle of the contract. The items will be sent to the CDE for three rounds of review. The first round will include the newly developed items prior to the ARP review. The second round will be the items that survived initial CDE and ARP review. These items will have all suggested CDE and ARP edits incorporated. The final round will be the items that reflect the second round of CDE edits. The items in the third review should be in final format and no further edits should be required at this stage. Note that there will only be two rounds of reviews for items for the History-Social Science CST assessments, since the ARP meetings are typically scheduled late in the summer. The CDE will review the items prior to the ARP meeting and then again once all ARP and CDE edits have been incorporated. The second CDE review will serve as the final review of items for History-Social Science.

Table 22. CDE Review of Newly Developed Items for 2006

Test	Content Area	Grades	First Review	Second Review	Third Review
CST	ELA	2, 3, 4	February 27 – March 10	May 22 – June 2	July 3 – July 14
		5, 6	March 6 – March 17	May 29 – June 9	July 10 – July 21
		7, 8	March 13 – March 24	June 5 – June 16	July 17 – July 28
		9, 10, 11	March 20 – March 31	June 12 – June 23	July 24 – August 4
	Math	2, 3, 4	March 6 – March 17	May 29 – June 9	July 10 – July 21
		5, 6, 7	March 13 – March 24	June 5 – June 16	July 17 – July 28
		Algebra I, Algebra II	March 20 – March 31	June 12 – June 23	July 24 – August 4
		Geometry, High School Math	March 27 – April 7	June 19 – June 30	July 31 – August 11
	Science	5, 8 *NCLB*	February 27 – March 10	May 15 – May 26	June 26 – July 7
		10 *NCLB*, Biology	March 6 – March 17	May 22 – June 2	July 3 – July 14
		Earth Science	March 13 – March 24	May 29 – June 9	July 10 – July 21
		Chemistry, Physics	March 13 – March 24	June 5 – June 16	July 17 – July 28
	History-Social Science	8	May 8 – May 19	July 10 – July 21	
		10	May 15 – May 26	July 17 – July 28	
		11	May 22 – June 2	July 24 – August 4	
STS	ELA	2, 3, 4	May 29 – June 9	July 31 – August 11	August 28 – September 8
		5, 6, 7	June 5 – June 16	August 7 – August 18	September 4 – September 15
	Math	2, 3, 4	May 29 – June 9	July 31 – August 11	August 28 – September 8
		5, 6, 7	June 5 – June 16	August 7 – August 18	September 4 – September 15
CMA	ELA	2, 3, 4, 5	May 1 – May 12	July 3 – July 14	August 14 – August 25
		6, 7, 8	May 8 – May 19	July 10 – July 21	August 21 – September 1
		9, 10, 11	May 15 – May 26	July 17 – July 28	August 28 – September 8

Test	Content Area	Grades	First Review	Second Review	Third Review
	Math	2, 3, 4, 5	May 1 – May 12	July 3 – July 14	August 14 – August 25
		6, 7, 8	May 8 – May 19	July 10 – July 21	August 21 – September 1
		9, 10, 11	May 15 – May 26	July 17 – July 28	August 28 – September 8
	Science	5, 8, 10	May 8 – May 19	July 10 – July 21	August 14 – August 25
CAPA	ALL	ALL	April 25 – May 5	June 19 – July 14	August 7 – August 18
STS grades 8 through 11 will be added to the review rounds in 2007, when development on those tests begins. CMA items for ELA and Math at Grades 4 through 11 will be a review of items selected from the CST banks for inclusion in the CMA.					

6. G. External Item and Task Reviews

The Item Development Coordinator will be in continual contact with the CDE and with ETS's content area Development Teams and item writers throughout the item-writing and editing process and will assure that ETS provides information to the CDE at all stages of the process. Once all internal ETS reviews are complete and all edits have been made, the items, along with all identifying information, will be prepared for review by the STAR committees, which include the ARPs and the SPAR panel. The tasks of the ARP will include the review of content (match to standard, technical quality) as well as review for bias and sensitivity issues (elimination of language, symbols, words, phrases, and content regarded as offensive to subgroups). The tasks of the SPAR panel include final review of all new items before they are included in field-test forms for compliance with *Education Code* Section 60614.

ETS will keep thorough and accurate notes throughout each committee review. Following each meeting, ETS assessment specialists and CDE staff will discuss notes from their item books so that both parties have the same copy of all edits, comments, concerns and suggestions from the committee.

Following the ARP review, and prior to the SPAR review, the ETS assessment specialists remove rejected items and incorporate committee changes to all remaining items. Both test development staff and editorial specialists conduct internal reviews prior to preparing items for each review stage. After all reviews have been completed, assessment specialists review this information before selecting items for field testing.

Assessment Review Panels

Committee membership will include educators and university and college subject matter specialists. In contacting potential committee members, ETS will assure a representation of gender and of the geographic regions and minority groups in California.

ARP members who are current school staff members will have the following qualifications:

- A bachelor's or master's degree in the subject area
- Three to five years of teaching experience in the subject area
- Demonstrated knowledge and experience with California content standards (this will be ascertained during the recruitment process)
- Special education credential (CAPA and CMA only)
- Experience with more than one type of disability (CAPA and CMA only)
- Three to five years as a teacher or school administrator with a special education credential (CAPA and CMA only)
- Bilingual and biliterate in Spanish (STS only)

SPAR Panel Review

The SPAR panel is convened by the CDE on the third Wednesday of each month. During the appropriate times, all new items will be presented in binders for SPAR panel review. The items will have been revised and proofread following the ARP meetings. Experienced ETS staff will be present at the SPAR meetings to facilitate the reviews.

6. G1. Description of and Augmented Timeline for Review Panels

ARP members will be trained to review newly developed items for relevance in terms of the purpose of the test, alignment with the California content standards, difficulty range, clarity, correctness of the answer and plausibility of the distractors. ARP meetings will also include a

review for issues of bias and sensitivity. In addition to the review of newly developed items, ARP members will also review versions of the proposed Released Test Questions (RTQs). Because these items are released to the public, it is important to have the ARPs review these items for the statistical characteristics of the items and match to the standards. After the newly developed items have been field tested, the ARP members are reconvened to review the resulting data from the field tests.

SPAR panel members are convened to review all newly developed items for compliance with *Education Code* Section 60614. No items will appear on STAR assessments without prior approval of the ARP, the SPAR, and the CDE.

ETS will increase the involvement of the ARPs in STAR test development and review activities. The following meetings will be added to the established ARP meetings:

Passage Review Meeting. STS ARPs will be assembled to review passages for the Standards-based Test in Spanish before items are developed to accomplish the following goals:

- Verification of proper distribution of commissioned vs. permissioned, type of passage (literary, functional, informational) according to the test specifications.
- Early opportunity to identify potentially biased or offensive content.
- Ability to edit commissioned passages to improve text or build in information to support item writing.
- Passages that reflect ARP direction for Spanish-language usage, punctuation, and grammar.

Members of the STS Reading-Language Arts ARP will meet to conduct passage review. Meetings will be held annually, early enough in the development cycle as to allow enough time for careful item development.

Blueprint Review Meeting. In keeping with the rollout schedule for the Standards-based Tests in Spanish, ETS will host an ARP meeting to review blueprints for Grades 5, 6, and 7 in February 2006. A review of the proposed blueprint will help guide item development, and assure an assessment that is parallel to the CST. A similar blueprint review will be conducted for Grades 8, 9, 10, and 11 in 2007.

Data Review Meeting. Once field testing for CMA and CAPA components has been completed, the respective ARPs will meet in August to review data. The panel will review and discuss items with questionable statistics, and the implications (if any) of instruction on student performance. These annual meetings will follow the same procedures used for ARP review of data in the CST and STS.

Forms Review Meeting. In keeping with the procedures developed for the CST, STS, and CAPA, ETS will also host a Forms Review meeting for the CMA ARP in September of each year. This important checkpoint in the test development process will provide the State with another opportunity to receive valuable feedback from California educators.

Released Test Questions Review Meeting. To maintain consistency across programs, ETS will host an ARP meeting to review released test questions for CAPA. This meeting will be held in conjunction with ARP data review meetings in August of each year.

Item Specifications, Test Specifications, and Sample Items Review Meetings. ETS will host a one-time meeting to review item specifications, test specifications, and sample items for STS, CMA, and CAPA. Such a meeting will help assure that ETS, the CDE, and

ARPs share the same vision for the content and format of the items for these new programs. These meetings will be held in April of each year.

CST Grade 3 ELA Meeting. ETS will host a meeting of the ad hoc group of ARP and Technical Advisory Group (TAG) members who had met in October 2005 in a continuing commitment to follow a three-year plan for improvement of this CST.

Table 23 outlines a proposed schedule for ARP and SPAR meetings for the 2006 development cycle. All dates are subject to approval by the CDE.

Table 23. Proposed Schedule of ARP and SPAR Meetings for 2006

2006 ARP and SPAR	
Meeting	**Date**
ARP — STS, Passage Review	February 7–9
ARP — STS, Blueprint Review (Grades 5-7)	February 21–23
ARP — CAPA, Item Specs, Test Specs, and Sample Item Review	April 4–6
ARP — STS, Item Specs, Test Specs, and Sample Item Review (Grades 5–7)	April 4–6
ARP — CST, Science (New Items)	April 10–12
ARP — CMA, Item Specs, Test Specs, and Sample Item Review	April 10–12
ARP — CST, ELA (New Items)	April 18–20
ARP — CST, Math (New Items)	April 25–27
ARP — CAPA, New Items/Tasks	May 16–17
SPAR — CSTs, Science, Math, ELA	June 21
ARP — CMA, New Items and Reporting Categories	June 21–22
ARP — CST, History-Social Science New Items, Data, and RTQs	June 27–29
ARP — STS, New Items	July 12–14
ARP — CST, Science Data and RTQs	July 18–20
ARP — CST, ELA Data and RTQs	July 25–27
ARP — CST, Math Data and RTQs	August 1–3
ARP — STS, Data and RTQs)	August 8–10
SPAR — CST, History-Social Science, CMA, CAPA, STS	August 16
ARP — CAPA, 2007 Operational Forms	September 1
ARP — CST, All Content Areas 2007 Operational Forms	September 8–9
ARP — STS, 2007 Operational Forms	September 29

6. G2. Review Panel Processes

ETS's assessment specialists will facilitate the ARP meetings.

Assessment specialists begin these meetings by conducting a training session on how to review items. ETS will also specifically train panel members in reviewing items for relevance in terms of the purpose of the test, alignment with the California content standards, difficulty range, clarity, correctness of the answer, and plausibility of the distractors.

ARP panels will also review items for bias and sensitivity issues. Below is a list of tasks commonly associated with bias and sensitivity review.

- Cultural diversity
- Diversity of background, cultural tradition, and viewpoints to be found in the test-taking populations
- Changing roles and attitudes toward various groups
- Role of language in setting and changing attitudes toward various groups
- Contributions of diverse groups (including ethnic and minority groups, individuals with disabilities, and women) to the history and culture of the United States and the achievements of individuals within these groups
- Avoidance of stereotyping and using language, symbols, words, phrases, or examples that are sexist, racist, or otherwise potentially offensive, inappropriate, or negative toward any group

After training, the first step of the item review process, panel members will be asked to review a set of items independently. As committee members individually review items for content validity, technical quality and problems of bias or sensitivity, they record their comments on the rating sheet. The next step in the review process would be for the group to discuss each item. At any time during the meeting, each content area group may divide further into smaller groups, depending upon the consensus of the group members and the numbers of items still to be reviewed. ETS staff will provide oversight of the smaller groups as well.

At data review meetings, the ARP will be asked to discuss items that have "poor" statistics. The committees will be asked whether there are instructional issues that have negatively affected the performance of the item or whether they have detected a content problem within the item. The committee will be asked to recommend whether to accept or reject each item for inclusion in the STAR item bank. The CDE will define the criteria for acceptable or unacceptable item statistics. The panel members will also use Differential Item Functioning (DIF) data to make judgments about the appropriateness or fairness of items to all individuals and subgroups. One or more TAG members will be invited to participate in ARP data review meetings.

SPAR Panel

The Statewide Pupil Assessment Review (SPAR) panel will assure that the test items conform to the requirements of *Education Code* Section 60614. ETS will assemble the items into field-test sets for the SPAR review to help keep the development process on schedule. If specific items are rejected by the SPAR panel, ETS will adjust the field-test sets as needed. The constructed-response writing tasks will also be prepared for review. If the SPAR panel rejects specific items or constructed-response writing tasks, they will be removed from the pool of available items. For the SPAR panel meeting, the set of items for review will be delivered to the CDE headquarters in advance of the meeting. The items will be accompanied by a cover memorandum describing the content areas and numbers and types of items being presented for review. The Test Development Project Lead or other designated staff will be available to respond to questions during the course of the meeting. Items will be picked up by ETS at the CDE headquarters at the conclusion of the SPAR meeting.

6. G3. Review Panel Training Materials

Two documents form the basis for the ARP review guidelines. The first is titled Guidelines for Item Writers, and the second is titled Guidelines for Bias and Sensitivity Review. ETS will work with the CDE to revise these documents as needed for the STAR Program. ETS will then use these documents for training at committee meetings as directed by the CDE. General guidelines for item review are listed below.

Guidelines for Reviewing Items for Content Issues

Does/Is the item —

- Have one clearly correct answer?
- Measure the intended California knowledge or skill?
- Test worthwhile concepts or information?
- Reflect good and current teaching practices?
- Reflect appropriate Principles of Universal Design?
- Present content that is free from bias against any person or group?
- Appropriate for English Learners (ELs)?
- Have a stem that gives the student a full sense of what the item is asking?
- Avoid unnecessary wordiness?
- Use response options that relate to the stem in the same way?
- Use response options that are plausible and reasonable?
- Avoid having one response option that is markedly different from the others?
- Use graphics or art pieces that are easily interpretable and directly related to the item?

Panelists will also need to make judgments on item stimuli based on whether or not each individual stimulus is —

- Required in order to answer the item
- Likely to be interesting to students
- Clearly and correctly labeled
- Complete, so that it contains the information needed to answer the item

ARPs will review newly developed items for fairness, as detailed in the ETS Standards for Quality and Fairness.

Guidelines for Reviewing Items for Bias and Sensitivity Issues

When reviewing items for bias and sensitivity problems, consider the following. Does the item —

- Contain language that is not commonly used across the state or has different connotations in different parts of the state?
- Discriminate in any way against individuals with disabilities?
- Have any references to religion?
- Reflect a gender or ethnic bias in any way?
- Make assumptions that all students are from the same socioeconomic group?
- Show anyone in a stereotypical manner?
- Have any offensive or demeaning words?
- Contain any controversial or emotionally charged subject matter?

When reviewing the set of items, consider the following. Does the set of items —

- Contain adequate coverage or representation of all subgroups?
- Contain artwork that adequately reflects the diversity of the student population taking the test?

ETS's assessment specialists will be responsible for preparing all relevant materials for the review meetings. These materials include the items with all associated coding, reference or resource materials, and guiding program documents, prepared and formatted in the style approved by the CDE. Item review materials will be organized in individual binders for meeting attendees based on guidance from the CDE. ETS will provide agendas, meeting handouts, and procedures for conducting the meetings to the CDE for review prior to the meetings.

In addition to the review binders and other meeting materials, ETS will prepare all necessary tracking sheets and other forms for committee members to record their judgments, and for staff to maintain systematic documentation of the proceedings and results.

6. G4. Documentation of Review Panel Meetings

ETS assessment specialists will be responsible for making sure all recommendations made by the committee members are recorded in a master item-review booklet. After the ARP meetings, ETS assessment specialists will meet with the CDE to discuss recommended changes to items. CDE will make all final decisions with regard to test items and accompanying materials. Through consultation with the CDE, ETS will make all CDE-suggested revisions to items and prepare the items for field testing. ETS's content area and editorial staff members will proofread the revised items to verify that the changes have been appropriately captured and that the items read as specified by the CDE. ETS will submit any revisions or replacements of items to the CDE for final approval and sign-off. ETS staff will summarize the committee meeting proceedings within the timeframe specified by the CDE, emphasizing in particular the committee recommendations for individual items. All meeting materials and summary reports will be archived by meeting date and committee name for the remainder of the contract period or as long as required by the CDE.

6. G5. Meeting Logistics

ETS Program management staff will be responsible for item-review meeting arrangements and associated costs, excluding costs for CDE staff, SBE members and SPAR panel members.

Specifically, ETS staff will:

- Contract for suitable meeting rooms that are well-lighted, have plenty of working table space and are comfortable for working groups
- Arrange for catering services for the meetings, including continental breakfast, morning and afternoon break refreshments, and lunch
- Arrange for meeting equipment for the work of the committees
- Submit names of committee members to the CDE for approval
- Contact potential review committee participants to announce meetings and to arrange for travel
- Reimburse committee participants per California State travel regulations for travel, lodging, and per diem
- Reimburse districts for substitute costs for teacher participants

All reviews will be held in Sacramento so that CDE and SBE staff may attend.

The CDE, SBE staff, and ARP members will receive ample advanced notification of meetings as well as copies of training materials for each meeting. In addition, committee members will receive materials and the necessary training at the beginning of each meeting. Committee members will also receive forms and directions on completing ETS's reimbursement process.

Recruiting New Panel Members

Item-review members will be recruited from California school districts, targeting classroom-level teachers and other educators. Recruiting letters will be sent to districts through both the superintendents' offices and the district STAR coordinators' offices. This dual approach not only provides important double-coverage, but also recognizes the vast range of California district sizes and administrative structures. Additional members will be sought from other stakeholder groups such as parents, school board members, higher-education professionals, and business and community leaders. ETS will strive for a broad representation from the diverse ethnic and geographic populations of California. Résumés of newly recruited reviewers will be submitted to the CDE for approval, giving the appropriate review time. ETS will be responsible for teacher substitute costs in accordance with the established district pay rate for substitute teachers.

Security of Materials for Review Meetings

All necessary measures will be taken to protect the security of meeting materials. Materials will be packaged and sent to the meeting site using a traceable and verifiable method of shipping. Once received at the meeting site, arrangements will be made in advance to keep the materials in a locked and secure area until ETS staff retrieves them in preparation for the meeting. Each binder and other set of secure materials or any other materials that the State deems secure will be individually numbered and identified to each meeting participant to make sure that all materials can be tracked.

Meeting participants will receive the secure materials only as needed per the agenda, will sign a confidentiality agreement, and will be instructed at the beginning of the meeting on how to maintain the security of these materials. At no time will participants be permitted to take the materials to another meeting room. When a room used by participants is vacated at any time, ETS will arrange to have the meeting room locked upon their departure to make sure that no unauthorized persons have access to the materials. All of the materials are collected at the end of each day of the meeting and kept secure overnight in a locked area inaccessible to unauthorized persons.

At the end of the meeting, ETS staff will gather, examine, and document all secure materials and account for each binder or set and its contents by number. Any missing materials will be reported immediately to the CDE. Secure materials that are no longer needed will be shredded onsite under ETS supervision or packed in a secure manner for traceable shipment back to ETS offices.

6. H. Field Testing

6. H1. Field Test Design

The analyses described below are general and would be applied to all field testing for the CST, CMA, STS and CAPA tests. In cases where the discussion only applies to a subset of the tests, this will be identified in the text.

Field-test items are presently positioned in the middle of the test in a block. ETS recommends that field-test items be dispersed through the operational test, so that field-test positions are varied, making it (1) easier to match their positions when the items are later used in operational forms and (2) more difficult to detect them as field-test items. Matching field-test position during test assembly could improve the stability of scores. For discrete items, items

could appear in every nth position in Year One and be shifted in the next year to deter recognition of the field-test items. When sets of items need to remain as an intact block, as is the case with ELA reading passages, different sets will appear in different positions when field tested. When items are used in operational forms, the field-test positions of the items will be maintained in the operational forms as closely as possible.

6. H2. Process for Analyzing Field Testing Results, Reporting Results to the CDE, and Presenting Analyses in the Technical Report

After the field-test data have been scored, all test items will be subjected to extensive statistical analyses. These analyses will show which items are at an appropriate difficulty level for the testing population and are free from any form of differential item difficulty for subgroups of the population. Additionally, ETS assessment specialists will conduct a confirmatory item to standard match for each of the content areas.

For all field-test items, the following sets of statistical analyses will be completed:

- Item analysis
- Differential Item Functioning (DIF) studies
- Calibration, scaling and equating

To assist psychometricians in overall evaluations of field-test items, the GENASYS system provides a comprehensive item evaluation module that incorporates information from a variety of analyses.

Differential Item Functioning (DIF). In most analyses of DIF, a focal group is identified that is the subject of the analyses; for example, Hispanic students. The rate at which this group answers an item correctly is compared to the rate at which the comparison group responds to the same item correctly. The analysis is carried out for groups who are at the same overall level of achievement. Consequently, if an item is measuring the same construct for both groups, the groups should perform similarly. If the item is differentially difficult for one group or the other, the item may be measuring something different from the intended content. DIF analyses will be carried out for all major subgroups of the testing population, using the Mantel-Haenszel (MH) statistic (see Dorans & Holland, 1993). Items that fail to meet the criteria for low differential item functioning will be revised for additional field testing or removed from the item pool.

Calibration, Scaling and Equating. ETS will utilize the proprietary version of the PARSCALE computer program to assure high-quality calibrations for the multiple-choice and constructed-response items.

For those content areas with embedded field-test sections, the field-test items will be calibrated simultaneously with the operational items and linked to the operational scale through a common-item equating design. The general process for calibrating and scaling the embedded field-test items will be similar for each administration.

The specific steps that will be completed for the calibration of items are:

1. Conduct preliminary item analyses for all the items, confirming that the statistical performance of all items is satisfactory. Items identified as being flawed or requiring revision (for example, negative item-total correlation) will be removed prior to calibrating the items.
2. Simultaneously calibrate the field-test items with the operational items, employing a one-parameter model for dichotomously-scored items and the Rasch partial-credit model for the ELA constructed-response items.

3. Evaluate model-data fit. ETS will evaluate fit statistics in conjunction with plots of model-data fit that are generated by the GENASYS system. Items flagged for potential misfit will be evaluated with respect to impact on test specifications, psychometric quality, and coverage of content standards and strands.
4. Link the item parameter estimates for the field-test items to the operational scale through a set of anchor or linking items consisting of operational items from previous administrations. Note that in standard one-parameter model analyses, this consists of setting the mean of the anchor item parameters based on the new calibrations equal to the mean of the anchor item parameters in the previous calibrations.

For tests that are newly field tested, the items will be calibrated simultaneously for each content area and a field-test scale created for form construction purposes. These items will be recalibrated following operational administration and placed onto the operational scale, which ETS will define in conjunction with the CDE. The item parameters will then be updated for scoring and reporting purposes. Once these tests become operational, the embedded field-test items will be calibrated simultaneously with the operational items.

Process for Analyzing and Reporting to CDE Data for all Subgroups

The field-test analyses will be analyzed and reported for all subgroups of the population as required by the *Education Code* and Title 5 of the California Code of Regulations. The analyses will be available for CDE review within eight weeks from the time of retrieval of the ELA, Mathematics, History-Social Science, and Science completed answer documents and test booklets from schools and/or districts.

Procedure for Providing Expert Advice on Accommodations and Modifications for Students with IEPs or Section 504 Plans

ETS's Office of Corporate Disability Policy will consult with the CDE on accommodations for students with individualized education plans (IEP) or Section 504 plans. ETS will make the résumés of staff experts available for CDE approval and will facilitate a conference call on any accommodations issues with the CDE.

6. I. Standard Setting

This section presents the general procedures that ETS employs in establishing performance standards for all tests aligned to California's SBE-adopted Content Standards. Those tests include the:

- CSTs — ELA Grades 2 to 11; Mathematics, Grades 2 through 11 (including subject-specific tests at high school), History-Social Science, and Science Grades 5, 8, 10 and subject-specific tests at high school
- CMA — ELA, Grades 2 to 11; Mathematics, Grades 2 to 7; and Science, Grades 5, 8, 10
- STS — ELA and Mathematics; Grades 2 through 11
- CAPA — ELA and Mathematics, Grades 2 through 11; and Science, Grades 5, 8, and 10

Standard setting will incorporate information from any required public hearing for newly introduced tests.

The purpose of the standard setting process is to define what a student must know and be able to do to meet the requirements of each of the defined performance levels. For each of these tests, the outcomes of the standard setting provide context for the placement of performance levels' cut scores by the SBE in consultation with the CDE for each of the

performance levels: Advanced, Proficient, Basic, Below Basic, and Far Below Basic. These results are based upon the review of test content, the use of student test data, and the knowledge and perspectives of teachers, curriculum specialists, school administrators, parents/guardians and community representatives.

Standards must comply with the federal *NCLB* requirements. In particular, California's standard defining Proficient is used to report the proportion of students reaching that level each year under the annual yearly progress requirement of *NCLB*.

All standard-setting activities will be led by an ETS standard setting expert who has experience in planning, facilitating, and documenting standard setting activities, training in special education and experience in developing and scoring alternate assessments, including training examiners and scorers.

6. I1. Procedure for Using Test Data to Identify Cut Scores

ETS will use item-ordered methods of standard setting for all content areas and grade levels of the CMA, the STS, and the CAPA. For the CMA and STS tests, ETS recommends using the Bookmark method. For the CAPA, ETS will employ a modification of the Bookmark procedure, called the Performance Profile Method, which has been used successfully for setting CAPA performance levels for ELA and Mathematics.

Training of Standard Setting Panelists. Training standard setting panelists thoroughly in the methodology is critical to producing meaningful and defensible cut scores. The majority of the first day is spent training panelists on:

- Standard setting (such as purpose, roles and responsibilities of facilitator and panel members, overview of the method and process, including an overview of the assessment)
- Content standards, which is facilitated by a assessment specialist
- Performance level descriptions
- Test items

An important component of training panelists on a standard setting method is giving them the opportunity to practice. In the training, ETS will include a brief overview of what a scale is and the implications of placing a cut point on that scale. After that overview, panelists will receive a detailed explanation of the standard setting method along with a chance to practice the method. For CAPA, representative examples of performance profiles may be used. ETS will provide panelists with an opportunity to practice making the judgments and then to discuss their judgments with other panelists. Only when they are comfortable that they understand the process and their role will ETS facilitators proceed. ETS will confirm their comfort with and understanding of the basic process with a formative evaluation form that each panelist will be asked to complete before continuing.

After the introduction to testing and standard setting, ETS staff will facilitate discussion with panelists to review the assessed content area as the primary driver for establishing performance levels. This will include a discussion of the connection between the content standards and the test specifications for each grade. For the CMA and STS, panelists will actually take the test in order to fully appreciate the cognitive requirements of each item. In addition, there will be discussion of performance level descriptions and labels associated with the performance levels. This discussion will attempt to elicit additional examples and clarifications of student knowledge and skills that are consistent with the performance levels.

In preparation for this discussion, as a pre-meeting homework assignment, ETS will ask the panelists to review the relevant content standards and to note (and bring to the meeting) three

to five indicators from the standards that they believe define each of the existing level descriptors.

6. I2. Process for Establishing a Panel Who Will Participate in the Standard Setting

ETS will work closely with the CDE and SBE staff and liaisons to identify panels of teachers, curriculum specialists, school administrators, parents/guardians, and community representatives to participate in the standard setting. ETS will recruit potential panelists from among names provided by the CDE, and the CDE and the SBE test liaison will have final approval of all panel membership. For each panel, the majority of panelists will be teachers currently teaching and currently licensed in the subject areas of the tests with not less than five years experience. The panels will be diverse in terms of geographic region and gender and will reflect the diversity of the State of California. For the CAPA and CMA, panelists also will have experience working with special education students. For the STS, panelists will also be bilingual and biliterate in Spanish and English. For the CAPA, the majority of the panelists will be trained and/or experienced in the administration of the CAPA.

CAPA and STS standard setting will be conducted concurrently during 2007; STS and CMA will be conducted concurrently during 2008. For each room, ETS will provide a lead facilitator, a content expert, a statistical/data entry expert, and a logistics/administrative specialist. For CAPA, the content expert will be an individual trained or experienced in special education. For CMA, every effort will be made to provide panelists who have special education expertise along with the content area background. A CDE and/or SBE representative will be in each room to provide a State-specific portion of the orientation.

Following the deliberations of the standard setting committees, the CDE will convene the TAG to review the recommended cut scores for reasonableness, consistency, and other characteristics and either suggest that the recommended cut scores be adopted by the SBE or propose adjustments prior to adoption.

6. I3. Procedures for Working with the CDE to Identify Potential Sites to Conduct Standard Setting Sessions and for Making Arrangements

ETS will be responsible for convening and hosting these standard setting meetings as described. ETS will work with the CDE to identify potential sites well-designed for such a meeting. Once ETS and the CDE have agreed on sites, ETS will be responsible for making the necessary arrangements for the meetings.

ETS will provide a continental breakfast, lunch, and coffee or drink service. ETS will cover all expense reimbursements to the panelists and all meeting facility and catering expenses and will assure that committee members are provided with reimbursement in a timely manner (two to four weeks). The proposed budget does not include CDE staff travel expenses.

ETS will provide a rehearsal of the standard setting workshops with the CDE and SBE in Sacramento. ETS will send two staff members, including the standard setting lead.

6. I4. Procedures for Developing Materials to be Used for the Standard Setting

ETS will consult with the CDE regarding the specific materials that will require CDE and SBE staff review and will develop a schedule for submission and review.

At a minimum, ETS will prepare and produce copies of the following materials:

- Biographical forms for prospective panelists
- Recruiting letters for prospective panelists
- Instruction letters for selected panelists
- Content Standards (supplied by the CDE)

- Meeting plans, agenda, and schedules
- Confidentiality and nondisclosure agreement
- MS-PowerPoint presentations used to train panelists in the standard setting method and specific procedures to be followed
- Practice tests and corresponding item maps to be used in training
- Practice cut score recording forms
- Actual test forms to be used in training
- Ordered item booklets and corresponding item maps for CMA and STS
- Ordered performance profile booklets and corresponding maps for CAPA
- Cut-score recording forms
- Evaluation forms — training, end of Round 1, and final

ETS will produce all materials in sufficient quantity for their intended use by panelists, ETS staff and the CDE and SBE representatives will arrange for shipping as required for standard setting meetings and planned reviews.

6. I5. Procedures for Producing a Technical Report

ETS will document the standard setting process and results and will present a written report to the CDE.

The data provided will include these essential results for each content area and grade level:

- Recommended cut score for each performance level
- Recommended cut scores plus and minus one and two standard errors of measurement and judgment
- Estimate of the proportion of students likely to be classified within each performance level (based on the recommended cut scores)
- Committee demographics

The report will also include the range of cut scores selected by the committee; frequency of selection of each cut score for each round (placement of the bookmarks); median, mean, and standard deviation for each round; selection of committee members; development of descriptors of the standards; training procedures; handling of intra- and inter-committee member differences; the role of impact data; and the handling of alignment of levels across grades and content areas as applicable.

ETS will also record any pertinent questions, observations, or statements made concerning how panelists made their judgments. These can be used in demonstrating the validity of the procedure and how the determination of final cut scores is reached.

To further support validity documentation, ETS will distribute short evaluation forms to panelists at the ends of Round 1 and the workshop. These evaluations will ask panelists to rate the degree to which they understood the process, materials and data, and what factors they considered in making their judgments. The results of the evaluation will demonstrate the degree to which panelists understood and applied the procedures in which they were trained.

6. J. Grade 3 Pattern of Results Recommendations

The STAR results from test administration years 2002–2005 showed improvement in performance in all grade/content areas except Grade 3 English Language Arts (ELA). This applies to mean scaled scores as well as percent proficient and above. After an extensive study of the Grade 3 ELA results, a task force including SBE representatives, the CDE, ETS, and members of the ELA ARP and TAG made several recommendations. These recommendations will be implemented in this scope of work. The following list of recommendations references the sections of this scope of work where the recommendations will be implemented:

Table 24. Summary of Recommendations

Recommendation	Scope of Work Section
• Make minor adjustments in item order and in number of passages to balance sessions in the Grade 3 ELA test for 2006.	7.A (Designing Test Forms)
• Continue working with the ARP to develop, field test, and use Grade 3 operational passages that are shorter and of less cognitive complexity.	6.G (External Item and Task Reviews)
• Schedule an annual ARP review of previous year's form to obtain recommendations on form construction for current test administration year.	6.G.1 (Augmented Timeline for Review Panels)
• Conduct additional research studies.	6.H2 (Analyzing Field Testing Results)
• Expanded communications strategy.	2.J (Communication with School Districts)
• Provide a more proactive outreach on what information is available about the tests and standards.	2.J (Communication with School Districts)
• Enhance the item release with more passage-related items, describing the linkage to standards and illustrating the variety of items associated with each passage.	2.H (Released Test Questions)

In addition, the following Three-Year plan will be followed to assure that the Grade 3 ELA CST changes to the desired form while preserving its year-to-year comparability:

Year 1 (2006 admin): Steps taken:

1. The number of reading passages was reduced by one (one less passage than in 2005).
2. The number of language passages was reduced by one (one less passage than in 2005).
3. The reading load (number of words) was reduced for 2006 to match the reading load in the 2004 test. That makes (in round numbers) 2,000 words at Grade 2, 4,000 words at Grade 3, and 6000 words at Grade 4.
4. Some test items were rearranged to give students some stand-alone items to start the first session (as there have been in Grade 2).
5. Test items were rearranged to balanced the reading load of each of the three sessions to a greater degree than in prior test administrations.
6. Items were grouped, whenever possible, by similarity of stems.

Year 2 (2007 admin): Steps to be taken:

1. With new item development in 2006, minimize the variety of item stems, using more consistent ways to ask the questions for each standard so that students have less "processing" to do to show what they know and can do.
2. Whenever possible, choose items for the 2007 operational forms that have consistent stems.
3. Order the 2007 operational items within a passage more consistently, and place general comprehension questions prior to vocabulary items.
4. Introduce new reading passages in the 2007 operational forms to be consistent with the instructional approach used in SBE-adopted materials. Replace at least two of the 2006 reading passages with new ones that were field tested in 2005 and 2006 (i.e., remove passages dating from 2001 to 2004). The impetus for this substitution is that the 2005 and 2006 reading passages reflect the ARP's 2004 directive to make the grade 3 passages reflect mid-year Grade 3 reading level, rather than end-of-year level.
5. Keep the number of reading passages in 2007 the same as in 2006.
6. Reduce the number of language passages in 2007 by one — that is, one fewer passage than in 2006 (using more stand-alone language items).
7. Keep the 2006 arrangement of giving students some stand-alone items to start the first session (as there are in Grade 2).
8. Continue to keep the reading load of each of the three sessions balanced as it was in 2006.
9. Put the review of the Grade 3 results as a specific part of the ELA ARP agenda for the July data review meetings, incorporated with the review of the 2007 draft operational forms; verify that the reading load has remained the same from 2006 to 2007. SBE staff and liaisons will be invited to these data review meetings.

Year 3 (2008 admin): Steps to be taken:

1. Field test the new items that minimize the variety of item stems, so that they can be incorporated in the test in 2009. Continue to develop new items that minimize the variety of item stems and arrangements of items associated with a passage.
2. Whenever possible, choose items for the 2008 operational form that have consistent stems.
3. Incorporate in 2008 one or two more reading passages that were field tested in 2005–2007 (i.e., remove passages dating from 2001 to 2004).
4. Reduce the number of reading passages in 2008 by one— that is, one fewer passage than in 2007 (using more items per passage).
5. Keep the number of language passages the same as in 2007.
6. Keep the 2006–07 arrangement of giving students some stand-alone items to start the first session (as there are in Grade 2).
7. Continue to keep the reading load of each of the three sessions balanced as it was in 2006–07.
8. Put the review of the Grade 3 results as a specific part of the agenda for the July data review meetings, incorporated with the review of the 2008 draft operational forms; verify that the reading load has remained the same from 2007 to 2008.

7. Component Task 7: Test Form, Test Booklet, and Answer Document Construction (CST, CAPA, CMA, STS, NRT)

7. A. Designing Test Forms for the CSTs, CAPA, CMA, and STS

ETS assessment specialists will build the individual field-test sets, reviewing items to avoid clueing. Once the operational test form has been built, but before it is shared for data review, the assessment specialists will review the operational form against the field-test sets and make changes in items, if necessary. As part of the CDE review process, ETS will provide a revised version of the field-test items when the operational forms for initial review are presented.

Forms for all STAR assessments will be evaluated for overall content, range of difficulty of items, and diversity of subject matter and approach, as well as multi-cultural and gender representation. The process described below details the steps taken toward these ends. The ability to build forms that incorporate these qualities begins, however, with the development of passages and items that, taken as a whole, produce an item bank that supports the creation of forms with appropriate depth and range.

ETS will take the following steps toward creating STS forms that will be equivalent in rigor to the CSTs. The first of these is to develop items for the STS to the same requirements as those that apply to CST items. To support this goal, ETS will:

- Have the same ETS Assessment Directors responsible for reviewing CST items also review STS items
- Train the STS ARPs in the same way as the CST ARPs
- Use the same Item Writer Guidelines/Test Specifications for the STS, except that sample items will be in Spanish
- Recruit for and facilitate ARP member overlap with members of the CST ARPs invited to serve on the STS ARPs
- Build forms of the STS within the same range of *p*-values as the CSTs

ETS will follow the principles of universal design in all test materials. While the application of these principles begins with the Academic Content Standards, they are especially important in materials that are used directly by students (i.e., test forms and answer documents). Section 7. D below discusses the application of Universal Design to Form Design and Production for Students with Disabilities.

7. A1. Content and Psychometric Criteria Used for Item and Task Selection and Ensuring Proper Test Form Assembly

The test construction schedule requires that the operational forms be built four to five months prior to their administration. This timeline allows the CDE at least 15 days to review the forms and the final proofs prior to printing.

During the operational test form development process consists of the following steps:

1. The Development Team receives the item statistics from the field tests. The content requirements of the test blueprint govern the item selection process for the operational forms. However, each operational item must meet stringent psychometric requirements as

well. Working with the CDE, ETS will verify the acceptable parameters for difficulty, item-total test correlation, and Differential Item Functioning (DIF) statistics.

2. Operational forms are built. Test development specialists identify an anchor set of items, which will comprise approximately 50 percent of the items in each content area. ETS will work with the CDE to identify anchor sets that meet both the content and psychometric guidelines for high-stakes tests.
3. The ETS Development Team will identify the remaining operational items. The full operational forms will be submitted to the CDE for approval.

ETS verifies that:

1. All items reflect the decisions of the ARPs and the CDE, and have been reproduced with 100 percent error-free accuracy.
2. The field-test sets are constructed to reflect a variety of strands and estimated difficulty levels.

All field-test sets will be submitted to the CDE for approval.

7. A2. Test Length and Composition and Test Alignment/Linking with Content Standards

Test Length and Composition

The tests will be constructed to accurately reflect the approved test blueprints. Because the blueprints identify the content standards to be assessed and the numbers of items to be tested under each content standard, they govern both test length and test content.

Having a large number of items covering each content standard is key to good test construction:

- It provides flexibility to item selection, as items can be "traded in" or "traded out" of the initial draft test form in order to meet the statistical requirements of the test.
- It is especially important for passage-based ELA items. Obtaining the appropriate test characteristics can be difficult if there are no extra items for a passage that can be substituted when needed. ETS will create an appropriate overage of usable items for every passage so that test construction can adhere to the test blueprints in every respect and reflect the highest standards of assessment.

The requirements of the STAR blueprints must be maintained, including the provision of retaining 50 percent of items, including the anchor item set, in the operational forms. The Mathematics field-test items must be placed with the other items in the Math strands. For ELA, the field-test items will be placed in various locations within the operational form.

7. A3. Item Selection for Choosing Items for Operational Forms

7. A3a. Process for Selecting Items/Tasks

Item Selection System for Test Development

For the item selection system in STAR test development, ETS will continue to use the STAR Item Bank System (see Section 5.). The item bank provides the test developer with a worksheet of items that are eligible to be selected for a new administration. Various statistical and classification information is available for these items, such as item number, item type, status, standard code, answer key, and statistical information (for example, point biserial correlation for the key and for each distractor, item *p*-value, and proportion of students choosing each distractor, IRT b-value).

Development of New Forms for Each Administration

When each STAR form is constructed, assessment specialists will choose and assemble a collection of test items designed to reflect a reliable, fair and valid measure of student learning within well-defined course content. Specific procedures and required test characteristics will vary from area to area. In constructing the test forms, the ETS assessment specialists will refer to the following content considerations on test form requirements.

Content Considerations:

Assessment specialists will adhere to all content requirements noted in the appropriate blueprints. The correct number of items will represent each standard and strand on each form.

Assessment specialists will select items so that the test content will be a thorough sampling of the knowledge and skills that are being measured. They will review the constructs and content included within each standard on the test blueprint and will make sure that items do not focus on a narrow range of the components within a standard.

Assessment specialists will review the comments of review committees and select items that follow this input.

For passage-based tests, the ELA assessment specialists will sequence passages to establish and maintain student interest in reading material. In order to accomplish this, test developers will:

- Make sure the passages are appropriately diverse in content and that the topics and writing styles presented are varied and appropriate to grade level
- Use a mix of previously published and commissioned passages
- Sequence passage types (literary, informational, etc.) to provide a varied reading experience
- Alternate passages by length

Assessment specialists will check that all the items selected have gone through the appropriate review procedures and have been approved for use in test forms.

Within each form, test developers will select items that reflect a balance of gender, ethnicity, regions, etc. A chart will be created that notes the distribution of names by gender, ethnicity, and type of role (traditional or non-traditional).

Items with and without artwork will be used in appropriate proportions and balanced within the test form. This will provide variation in appearance of page spreads and in the testing experience.

Assessment specialists will review individual items for dated content. Items with content that has become dated will not be used.

Items will be carefully reviewed for instances of clueing - information in one item or stimulus providing information that clues the correct answer to another item.

Statistical Considerations:

Before the ETS assessment specialists begin to build a live test form, they are trained in reading, interpreting, and using the item statistics that appear on the data label that will accompany each item.

Most of the individual items selected by the assessment specialists will have *p*-values (a measure of difficulty) that range from 0.35 to 0.80. Some items may be chosen outside of

this range in order to provide more meaningful and accurate scores for students at a wider range of performance levels.

Assessment specialists will choose a collection of items yielding overall difficulty that falls within a specified range. **This means that each parallel form will be built with a comparable mean difficulty for each strand or reporting category and for the total test.** Also as a result, forms in each content area will be equivalent in difficulty to the forms in the same content area from year to year.

Items will be chosen in which all four answer options are functioning (attracting some students). Both p-values and point-biserials will be examined for each distractor.

The differential item functioning (DIF) flags and DIF statistics will be carefully considered.

Items used on any form will have item-total correlations above 0.20.

Arrangement of Items:

Equating anchor items will be placed in each form first, and the sequence of the equating anchor items will be kept consistent from form to form.

The initial items on a form and in each session will be relatively easier than those items that follow so that many students experience success early in each testing session.

The remaining items will be sequenced within a form and within a session by alternating easier and more difficult items.

The distribution of correct-answer positions will be checked so that there will be an approximately equal number of correct answers in each of the four positions.

Test developers will not sequence the correct answers into any sort of pattern.

ETS will attend carefully to the flow of item content within the form; the area of knowledge students will be required to use to respond to an item will not be extremely different from the previous item.

Field-Test Items:

Assessment specialists will choose items to be field tested on each form according to the specifications and the item development plan for each test. These items will be chosen from items that have successfully passed the complete review process. All field-test items will be embedded in fixed locations throughout the test in a given year, rather than appearing as a group at the end. They will appear in the same positions on every version of the form. The locations of the field-test items may change from year to year.

Versions of a form will differ only with regard to the embedded field-test items in a given year.

ETS will submit all field-test sets to the CDE and will make changes and any necessary substitutions as requested by the CDE.

Field-test items will be selected to assure that the content requirements noted in the test blueprint and the item utilization plan will be met in the future.

Page Layout:

Tests must be readable and inviting and have ample white space. Items must be positioned carefully on each page so that their sequence is clear to all students. Artwork and graphics are important in tests as a way to lend interest (when accompanying a passage, for example), as stimuli (for example, for an algebra problem), and benefit the visual learner. All artwork and graphics must be clearly drawn and clearly reproduced. It is

also important that all artwork is easily accessible to visually impaired test takers and English language learners. ETS designers will avoid using intricate fonts, small type sizes, gray shadings, and other complexities that might tend to disadvantage some groups of students. For the STAR assessments, ETS will continue using the layout developed for previous administrations.

Items connected to a common stimulus or passage will be, whenever possible, placed on the same page as the stimulus or on a facing page. This is not always possible in ELA because of the length of passages and the number of items that are associated with a passage.

Test developers will carefully consider how items will appear on the printed test pages. They will work with the editorial and publishing staff to create pages that look inviting but have limited wasted space. Because the field-test items will always be in the same location within the different versions, some versions may have more white space because of differences in passage or item length within the field-test items.

ETS will deliver to the CDE a form planner for each form that lists the linking items, the code that identifies the strand for which each item will be used to generate sub scores, the item key, the item statistics from the previous administration, the proportion of students responding to each answer option, the item position on the proposed form, and summary data for the total test, e.g., mean p-values, b-values, and point-biserials, etc. The form planner will also be provided in electronic format (readable by the STAR item bank) at the same time as hard copy delivery.

During the composition of the pages for the operational forms, the field-test items will be incorporated into the forms to create the required number of versions. The layout of each version will be carefully reviewed for completeness of the items, clarity of the artwork, and effective use of white space. ETS Assessment Specialists and editors will review each composed page. Furthermore, the assessment specialists will verify that each item is correct and that all parts of the item are present. The editors will do a technical check that will include verifying accuracy of headers, footers, folios, sequence numbers, and directions. They will also perform an end-line check against the original version of the item.

The CDE will have at least 15 working days for review and approval prior to printing deadlines. The CDE will have at least 15 working days for review and approval prior to printing deadlines.

7. A3b. Rotation of Standards Coverage on the CSTs

The item development plan in Section 1 takes the needs for rotation of standards coverage into account, developing items in proportion to the need to maintain an adequate supply of items in the item bank for each standard. The Assessment Director for each subject area will review each year's plan and make sure that the numbers of items to be developed to support the assessment of each standard, tested each year, that the number of items tested each year remain the same per content standard, and that the standards identified for testing on the CSTs are assessed. The plan will maintain a high degree of accuracy in the longitudinal test results across forms of the tests.

7. A3c. Year-to-Year Comparability of Scaled Scores and Performance Levels

Obtaining score comparability from year to year is the result of many processes throughout the test construction, administration, calibration and equating process. In addition to the anchor set model, ETS will employ other practices recommended in the literature for facilitating form-to-form comparability, such as:

- Choosing and employing anchor-item sets so that the anchor items in each new form are proportionally representative of the full test with regard to content specifications and test difficulty
- Scaling and linking all field-test items to the existing IRT scales
- Assembling tests to exacting content and IRT-based statistical specifications at both the cluster and total test levels, so that the test forms are as parallel as possible and greater stability of the equating relationship is achieved, resulting in greater stability of the scaled scores from year to year
- Assembling tests by considering model-data fit and selecting items that best fit the one-parameter model (subject to meeting content specifications)
- Rotating anchor sets to minimize item exposure and using non-anchor items to independently check the consistency of linkings based on the anchor sets in post-equated designs
- Linking items will appear in the same locations in the new and reference forms of the test so that position effects are controlled, which results in a more stable and unbiased equating relationship
- Carrying out test administrations in standardized ways so that students are treated with equity and common standards may be applied
- Making equating samples sufficiently large to estimate the item parameter estimates (and score) statistics to achieve greater scale stability

7. A3d. Procedures for Linking and Equating Test Forms

Equating tests is essential in order to provide scaled scores that have similar meaning from year to year. The CST and CAPA as well as the CMA and STS tests will be equated using an anchor-test design where a block of the same items is embedded in adjacent year tests. A new anchor block is selected each year from the previous year's test. The anchor block provides a means of placing all of the calibrated, but as yet unscaled new form items on the reference (base) form scale. ETS uses the PARSCALE statistical software to equate.

The new form items are transformed to the reference IRT scale using the Stocking and Lord (1983) test characteristic curve procedure. When all the items for both the new and old forms of the test are on scale together, an equating procedure (described in more detail in Section 11.D.3) is used to find the number-correct score on the reference form of the test that corresponds to the number-correct score on the new form of the test.

Through this process, the scaled score or number-correct relationship on the reference form is transferred to the new form of the test. The result is an equated scoring (conversion) table that can be used to convert the number-correct scores on the new form of the test to the appropriate (equated) scaled score.

While this model is very strong, some variability is introduced through the use of chained anchors (i.e., blocks of items that are common for adjacent years but not across several years). Were the same set of anchors used for longer periods of time, greater score stability might be achieved. Haertel (2004) has found that, at the aggregate level, item sampling appeared to play a large role in the variability of scores. This would be a way of addressing this source of score variability. The danger in repeating the same anchor over time is that these items may become over-exposed and change their characteristics which may result in scale drift.

7. A3e. Process for Providing Recommended Item Selections

For the first review of each operational form, ETS will provide a form planner with all required data for every item, one item per page item cards with the data for that item, and hard copy mock-ups of the proposed form in item order. The CDE consultant will then be able to review the proposed form with the STAR item bank available, should any substitutions be desired. The schedule in Section 1 details the sequence and duration of each of these reviews. By the end of each review period, the CDE content consultant will have the opportunity to share their feedback on the proposed form with their ETS assessment specialist counterpart.

7. A3f. Plan to Ensure that Tests will Include Test Items/Tasks of Differing Levels of Difficulty

The STAR tests are built to statistical targets in the form or a test characteristic curve and the conditional standard error or measurement (CSEM) curve. The assembly targets were developed using items from test forms that had been assembled to blueprints as well as using overall reliability and CSEM requirements. The equating blocks are assembled to proportionally adjusted targets. When these targets are met, it assures a distribution of item difficulty across the test as a whole. However, it does not guarantee that each test form will have the same variation of item difficulty within a strand or reporting cluster. Targets will be developed for smaller units of assembly than the test to help improve that parallelism of the test forms. While this will place greater demands on the item pools, it will yield greater parallelism between test forms and increased stability in the scores.

7. A4. Procedures for Providing the CDE with Test Forms that Include Proposed Items and Item Statistics for Review and Approval

In addition to the steps described above, ETS has scheduled several reviews subsequent to agreement on the proposed test forms. The second review of the proposed form will include the version of the operational test that shows the changes specified by the CDE content consultant in the first review. As in the first review, field-test items that also reflect changes following the first review of those items will be provided for every version of the form.

The third review will consist of final PDF versions of forms, both the operational form and the items for each field-test version. The field-test items will not be merged with the operational form so that CDE is not required to print the complete form of each version. When this form is presented to the CDE for review, the CDE will also receive a list of the changes requested following the second review so that the CDE content consultant does not have to review the entire batch of operational items and field-test items.

Each review of the form will be accompanied by a revised electronic and hard copy form planner.

7. B. Constructing Test Booklets for the CSTs, NRT, CAPA, CMA, and STS.

The following scannable test booklets will be produced for the STAR Program:

- Grade 2 CST test booklet
- Grade 3 CST test booklet
- Grade 3 NRT test booklet
- Grade 2 STS test booklet
- Grade 3 STS test booklet
- Grade 2 CMA test booklet
- Grade 3 CMA test booklet

The test booklets will provide fields to capture student responses and demographic information; hence the test booklet and answer document are one and the same for these tests.

Test Book Development Process

The development of test booklets is a collaborative process among ETS, operations subcontractor, Pearson Educational Measurement (PEM) and the CDE. Test booklets will be produced according to the following standard production process. The CDE retains final review and approval for testing materials prior to printing. The process includes the following steps:

1. Print File to PEM
2. Page Proofs
3. Page Proof Revisions
4. Second Page Proof Revisions
5. Production Phase

Print File to PEM

Early each October, ETS will provide PEM with print files for all scannable test booklets. In addition, ETS will provide bookmaps for each test booklet, and will specify which pages are common or variable across test book versions.

Page Proofs

Once PEM receives the print file from ETS, the following steps occur as part of the page proof process:

1. Demographic pages are added to the test booklet and the appropriate variable and common pages are put together to create the test booklet.
2. PEM then sends the document to their Proofreading Department with the standard proofing checklist attached. As part of this process, the directions in the test booklet are checked for format, grammatical correctness and proper use of conventions.
3. Proofreaders review the document using the bookmaps and item matrices for comparison and mark any corrections or revisions that need to be made.
4. The PEM Forms Project Director reviews the revisions and makes any additional required edits.
5. The PEM Forms Designer updates the draft with all marked revisions.
6. The process of reviewing and revising is continued until the document is determined to be “clean and final.”
7. After the review, the PEM Forms Designer creates the digital file complete with items and embedded graphics.
8. Page proofs of the test booklet will be printed and delivered to the CDE for its review. Project schedules include five business days for the CDE to complete a review, make revisions on the page proofs, and return the proofs to PEM to complete the workflow process.

Page Proof Revisions

When PEM receives the page proofs, their design team incorporates the necessary changes to the test booklet. PEM then sends the test booklet to their proofreaders, who compare it to the page proofs received from the CDE. If necessary, the Forms Designer makes further changes and the test booklet is then returned to the proofreaders.

Second Page Proof Revisions

After the page proof revisions are complete and verified by PEM's Proofreading Department, they will be delivered to the CDE for final review and approval.

During this process, the PEM Forms Department will continue to provide page proofs to the CDE until the final test booklet design is approved. Upon approval, the Forms Designer will create a digital copy, check it against the page proof, and send it to the printers. The document will go through a final verification process when the printer produces and sends the blueline of the document to the Forms Department, where accuracy and print image integrity will be verified.

Production Phase

A document will be released and declared "clean" only after all review criteria have been satisfied. Only then will PEM's printing division receive approval to print.

PEM uses the following detailed production processes to produce quality documents:

1. Press sheets are checked during the print run to verify thorough ink coverage, color and overall print quality.
2. Operators select sample documents for color proofing and quality testing at predetermined locations throughout the print run for testing.
3. Forms are carefully inspected for squareness of cut, exact positioning of tracks, codes, text and response positions, and the quality of printing.
4. Forms are released for subsequent activities only after they meet quality control standards.
5. After releasing the documents, the printed signatures move to the bindery area for binding, wrapping, and packaging in accordance with project specifications.
6. During binding, bindery personnel establish in-process checks to further verify quality control.
7. A final internal review is performed before documents are shipped to customers.

The test books for Grades 4–11 will be printed by ETS and willl be non-scannable. The page proof process will be similar as with PEM with the CDE receiving final laser for review and signoff, refer to the schedule given in Section 6 for these reviews.

7. C. Designing and Constructing Answer Documents for the CSTs, NRT, CAPA, CMA, and STS

PEM will design and print answer documents to correspond to the test booklets. The school and grade ID sheets will identify the school code, grade, teacher name, and other information needed for aggregating student data. The demographic intake sheet will collect demographic information for non-tested students.

Depending on the specific requirements dictated by grade and subject area, the scannable answer documents are single sheets, four-page folders, 12-page answer booklets, or self-contained test booklets (the Grade 2 and Grade 3 test booklets as described in Section 7. B). Answer documents will allow space for collecting responses to multiple-choice or the writing prompt. All answer documents can be pre-coded by either printing the specific information directly onto the document or by applying a machine-readable label containing the appropriate information.

The quantities of answer documents, school and grade ID sheets, and demographic intake sheets required for the STAR Program are provided on the following pages in Table 25. The

figures provided are based on the testing volumes included on page 9 of the STAR RFS and historical print overages required for this program.

Table 25. Printing Quantities for Answer Documents, School, Grade ID Sheets, and Demographic Intake Sheets

Document	No. of Pages	2007 Print Quantity	2008 Print Quantity	2009 Print Quantity
Grade 4 Answer Folder	4	665,000	665,000	665,000
Grade 5 Answer Folder	4	640,000	640,000	640,000
Grade 6 Answer Folder	4	630,000	630,000	630,000
Grade 7 Answer Folder	4	610,000	610,000	610,000
Grade 8 Answer Folder	4	620,000	620,000	620,000
Grade 9 Answer Folder	4	685,000	685,000	685,000
Grade 10 Answer Folder	4	660,000	660,000	660,000
Grade 11 Answer Folder	4	610,000	610,000	610,000
Grade 4 March Writing	12	565,000	565,000	565,000
Grade 7 March Writing	12	520,000	520,000	520,000
Grade 4 May Writing	12	100,000	100,000	100,000
Grade 7 May Writing	12	90,000	90,000	90,000
Math End-of-Course Answer Sheet	2	1,941,444	1,941,444	1,941,444
Science End-of-Course Answer Sheet	2	1,080,980	1,080,980	1,080,980
History-Social Science End-of-Course Answer Sheet	2	1,575,966	1,575,966	1,575,966
CAPA Answer Folder	4	60,000	60,000	60,000
Grade 4 CMA Answer Folder	4	12,000	12,000	12,000
Grade 5 CMA Answer Folder	4	12,100	12,100	12,100
Grade 6 CMA Answer Folder	4	12,200	12,200	12,200
Grade 7 CMA Answer Folder	4	12,500	12,500	12,500
Grade 8 CMA Answer Folder	4	12,600	12,600	12,600
Grades 9, 11 CMA Answer Folder	4	26,500	26,500	26,500
Grade 10 CMA Answer Folder	4	13,200	13,200	13,200
Grade 4 STS Answer Folder	4	18,000	17,900	18,300
Grade 5 STS Answer Folder	4	12,100	12,100	12,000
Grade 6 STS Answer Folder	4	6,100	6,100	6,100
Grade 7 STS Answer Folder	4	6,300	6,100	6,200
Grade 8 STS Answer Folder	4	NA	6,300	6,200
Grade 9 STS Answer Folder	4	NA	7,100	7,100
Grade 10 STS Answer Folder	4	NA	6,700	6,700
Grade 11 STS Answer Folder	4	NA	6,200	6,200

Document	No. of Pages	2007 Print Quantity	2008 Print Quantity	2009 Print Quantity
Math End-of-Course STS Answer Sheet	2	NA	28,800	24,850
Demographic Intake Sheet	2	90,000	90,000	90,000
School and Grade ID Sheet (SGID) (headers)	2	571,700	572,700	572,400

PEM's Forms Department will provide the CDE and ETS with a mockup design layout of the scannable documents before any composition begins. This will allow all parties an opportunity to make major changes to the forms without affecting schedules and without incurring significant costs.

Once agreement has been reached on the general layout of the documents, the materials will be composed in accordance with the determined specifications. The use of specialized forms design software minimizes the effort involved in traditional methods of form development and, in the event that changes are required, reduces the time required to generate revised proofs.

After each form is created, PEM staff will thoroughly inspect each document to verify that it matches the text and specifications received from the CDE and ETS. When all corrections are made, proofs will be sent to the CDE for review. A duplicate copy of each proof will be sent to PEM's software development team for technical review of scannability, oval placement, and spine code assignment.

Once the proofs are approved by the CDE and by PEM's software development team, the printing plates will be made and the documents will be printed at one of PEM's printing facilities.

Immediately after printing, sample documents are selected for testing from predetermined locations throughout the print run. These forms are carefully checked for accuracy of cut, exact positioning of timing tracks, codes, text and response positions, and the quality of printing. Forms are released for subsequent activities only after they have met all quality standards.

Answer documents may be revised annually to accommodate information necessary for the program. PEM will work with ETS and the CDE to develop and revise answer documents.

Single Record for Each Student

Answer documents will be designed to produce a single complete record for each student, which will include demographic data and all test scores for that student. Table 26 identifies the type of answer documents that will be utilized at each grade level.

Table 26. Answer Documents Utilized by Grade Level

Grade	CST	NRT	Writing	Math EOC	Science EOC	World History EOC
2	✓					
3	✓	✓				
4	✓		✓			
5	✓					
6	✓					
7	✓ (includes NRT)		✓			
8	✓			✓*		
9	✓			✓*	✓*	✓*

Grade	CST	NRT	Writing	Math EOC	Science EOC	World History EOC
10	✓			✓*	✓*	✓*
11	✓			✓*	✓*	✓*
✓* Students at this grade level may take an end-of-course test.						

For those students in Grades 2, 5, 6, and 8, demographic information and responses will be captured on a single, multi-page document. Students in Grades 3, 4, 7, 9, 10, and 11 will be assessed using multiple-answer documents as shown in Table 26.

All STAR answer documents will contain uniquely numbered lithocodes that are both scannable and eye-readable. The lithocodes will allow all pages of the document to be linked throughout processing, even after the documents have been separated into single sheets for scanning.

For those students testing on more than one answer document, lithocodes will link their demographics and responses within a document while matching criteria will be used to create a single record for all of the students' documents. PEM will use the matching criteria listed below.

- **Grade 4 CST and Grade 4 Writing | Grade 7 CST/NRT and Grade 7 Writing**
  1. SSID (Statewide Student Identifier, formerly CSIS) number
  2. First name, last name, date of birth, gender
- **Grade 3 CST and NRT**
  1. Pre-ID Barcode number
  2. SSID number
  3. First name, last name, date of birth, gender
- **Grade level (8, 9, 10, or 11) and end-of-course tests (Math, Science, and/or World History)**
  1. Pre-ID Barcode number
  2. SSID number
  3. First name, last name, date of birth, gender
- **CAPA.** Twenty percent of the CAPA tests are scored more than once (two answer documents received). The answer document indicates whether the test was scored by the examiner or the observer. The two answer documents will come in under the same school and grade ID sheets and need to be matched.
  1. First — SSID ID number
  2. Second — First name, last name, date of birth, gender
- **Data on STS** grade-level answer documents and STS Math end-of-course documents will be matched using the Pre-ID barcode number, SSID number, and the student's first name, last name, date of birth, and gender. No matching is required for CMA as all student responses for this assessment are captured on a single answer document.

End-of-course Test Answer Documents

Separate answer documents for CST Math, Science, and World History will be provided. Because students taking these end-of-course subject tests will also be testing on a grade level

answer document, full demographics are not required on the end-of-course document. Only those demographics required for matching the student's end-of-course scores to their CST grade-level scores (Pre-ID barcode, SSID, first name, last name, date of birth, and gender) are required. As a result, a two-page (single sheet) answer document will be used for Math, Science, and World History. Space will be provided for the tester to indicate which end-of-course test (for example, Algebra I) he or she is taking.

With the addition of end-of-course answer sheets to the STAR Program, matching for students at the higher levels will need to take place on up to four answer documents (grade level, Math, Science, and World History). The use of Pre-ID will greatly reduce the potential incidence of non-matching records, but non-matches may well occur, as currently is the case with the Grade 4 and 7 CST and Writing assessments, and the Grade 3 CST and NRT. In these instances, all student data and results are reported, but as individual records. The Pre-ID record layout will be modified to include fields for the end-of-course tests, so that end-of-course answer documents can be appropriately pre-identified.

A separate STS end-of-course math answer document will also be provided for the 2008 STS field test and 2009 operational administration. PEM will match STS grade level data with STS end-of-course math data to create student records.

Producing Pre-ID Answer Documents

Districts will have the option to pre-identify student answer documents for each of the test types (CST/NRT, Writing, CAPA, CMA, and STS). The following is a description of the process that will be used.

1. ETS will transmit Pre-ID files and enrollment orders for Pre-ID documents to PEM 42 calendar days prior to the first test date. The enrollment order will indicate whether the district chooses Pre-ID documents or labels.
2. Pre-ID files will be sent to PEM by administration and test type (CST/NRT, Writing, CAPA, CMA, and STS).
3. ETS will transmit Pre-ID files and enrollment orders to PEM 21 calendar days prior to the first test date.

Depending on the method of pre-identification selected, districts will receive their pre-identified student answer documents with their other testing materials, including additional non-pre-identified answer documents (10 percent overage for every school and five percent for every district) for those students not identified on a Pre-ID file. As part of a separate shipment, districts will receive any pre-identified student labels ordered.

Information not provided in a Pre-ID file will be hand-coded at the time of testing. Full demographics will reside on the CST multiple-choice grade-level answer documents, CAPA, STS, and CMA answer documents. Only those demographic categories required for matching will reside on the CST Writing answer document, end-of-course answer documents, and the NRT answer documents, thereby reducing the burden of gridding on the districts.

7. D. Forms Design and Production for Students with Disabilities for the CSTs, NRT, CAPA, CMA, and STS

Refer to the accompanying Schedule for Project Deliverables and Activities in Section 1. A plan for specific details about the design and production of forms for students with disabilities, as well as for review and approval by the CDE.

Universal Design Principles

Tests including operational items, field-test items, and test-bank items will be developed and administered in a manner that maximizes participation of students with special needs and allows for accommodations to the extent required by law, in accordance with the Individuals with Disabilities Education Act, 20 U.S.C §1400 et seq. (IDEA).

The goals of Universal Design are to increase the number of students with special needs who are able to take the STAR Program without special accommodations and to make accommodations easier to implement.

The specific steps needed to continue implementing Universal Design principles for the STAR Program include:

- Involving special needs stakeholders in test design and review processes
- Examining items for evidence of disability bias and eliminating any such items
- Using "plain language" in items
- Adjusting font size for ease of reading
- Minimizing the use of italic typeface
- Using highly distinguishable symbols on graphics

These principles will be applied consistently in the development of the STS and CMA assessments as they have been in CST and CAPA development.

7. D1. Braille Version of all Test Forms and Answer Documents

ETS test development, editorial and production staff will implement the requirements of the CST, CAPA, CMA and STS assessments regarding braille tactile graphic materials and materials for the other special versions, including the following guidelines for braille tactile graphic materials:

- ETS test developers will create all items so that graphic material does not contain clues or omit necessary information, nor will the graphics contain material that is unnecessary for responding to the prompt.
- ETS will review the descriptions of graphics and illustrations to maintain concise and meaningful information. These descriptions will be included as transcriber's notes throughout the test and will also be included in the Teacher's notes to the braille edition.
- ETS will practice the protocol of placing keys or legends that supplement reading graphics at the top left of the tactile graphic or on the left-hand facing page.
- ETS will use horizontal braille labeling on graphics.
- The ETS assessment specialists and special versions editor will maintain charts and graphs on one page when possible.

Test developers will also be involved in the process of developing the accommodated versions.

To make a determination as to the appropriateness of items for visually challenged students, ETS will seek the recommendation of those individuals experienced in brailling. First, test developers will provide the test form to the braille vendor for review.

The brailler's recommendations to the test developer may consist of:

- Identifying items that should not be brailled due to the inappropriateness of the content for a visually impaired student

- Suggestions for the scripting of art or graphics related to an item
- Suggestions for minor modifications to art or graphics for braille reproduction

The braille version will use uncontracted braille for grade 2 and will use contracted braille for grades 3 through 11 and end-of-course assessments. The *CAPA Braille Examiner's Manual* will be produced in contracted braille.

Contracted braille is a combination of braille letters and short-form words. Uncontracted braille will only be used if any portion of the test requires the test taker to spell words.

The test developer will work with the brailler to make final recommendations for a braille test form. Special consideration must be given to the number of items recommended for omission. Test development and psychometric staff will review the impact on the test specifications and score reporting requirements for subtest scores.

ETS will provide braille test forms for the STAR administrations. One new form will be provided for each yearly administration. All braille books will exclude field-test items and will be individually wrapped and labeled. However, these will be included with the shipment of other test materials to the schools and districts.

The quantities of the braille-print test books for each administration will be estimated after surveying STAR LEA district coordinators and the administrators of Special Education Local Plan Areas (SELPA) in order to accurately identify the numbers of tests to be ordered for each test administration. Actual quantities to be printed will be based on school and district orders obtained during the ordering process for each administration.

ETS will produce sufficient quantities of braille test books and supporting answer documents to support the initial orders derived from the process above, any supplemental orders and those necessary to support any review processes by the CDE and its staff.

7. D2. Procedures for Producing a CD-ROM Version and a Large-Print Version of the Test

CD-ROM and Large-Print Versions

Large-Print Version

For large-print forms, test developers will identify those items that must be particularly attended to by the staff responsible for producing the large-print forms. Test developers will provide directions to production staff, relaying information about Mathematics items that involve measurements that cannot be enlarged, about art or graphics that may be affected by enlargement, and about the spacing of materials that affects performance on items. The goal is to avoid introducing factors that affect item performance, thus maintaining the validity of all items. The large-print version will be produced in a 20-pt. Arial font format.

As an application of Universal Design Principles, the large-print edition is not just a larger edition of the operational form, but is adjusted to accommodate the needs of students. A template will be created that adjusts the amount of white space, leading, and other factors that impact the accessibility of the form for students.

As with the braille forms, ETS will gather initial order quantities from LEA coordinators. ETS will produce quantities sufficient to meet those initial orders, estimated supplemental orders based upon past experience, and a small amount to satisfy CDE review of the documents.

CD-ROM Version

ETS will provide CD-ROMs for each test administration to accommodate students with an IEP or Section 504 plans.

CD-ROM versions will be produced for each operational content area and assessment type for the administration each year. CD-ROMs will be capable of variable font characteristics (size, color) and background color, as well as the inclusion of an audio version of the test. ETS has planned for the production of the text component in HTML format.

The audio portion can be accommodated in one of two alternatives. There are new software technologies available that voice-enable HTML or PDF formats which have proved useful for those with mild visual impairments. The other alternative is to pursue a more traditional approach that utilizes a straightforward audio version of the test on a separate CD-ROM. ETS will work with the CDE to determine the most efficient solution to audio delivery and implement the one that best meets the needs of this segment of the testing population.

The CD-ROMs will be produced in quantities to meet the demands of the program as canvassed through the order management system.

ETS will also print test books to accompany CD-ROMs for students who want to follow along with an actual test book. These test books will be sealed and distributed to the districts and schools along with the CD-ROMs, upon request. Actual quantities to be produced will be based on school and district orders obtained during the ordering process with the overages (plus ten percent for schools and plus five percent for LEAs) applied for each administration. ETS will produce sufficient copies of the CD-ROM and large-print versions and will distribute them to the LEAs upon their request.

7. D3. Procedures for Producing Test Coordinator Instructions and Test Examiner Directions

ETS will provide detailed district STAR coordinator instructions and examiner directions to support the test for each of the special test versions — braille, large-print, and CD-ROM. Included in all manuals will be clear and concise directions for administering the tests. ETS will continue its current practice of producing a single coordinators manual and examiners manual that encompasses both the standard and special versions.

Instructions in the manuals for use in administering braille editions of the test to students will mirror the standard administration directions as much as possible. However, they will reflect the changes made to the non-inclusion of the field-test items, as well as any special instructions for administration specific to the braille version of the assessment.

The directions for administration specific to the large-print edition will be similar to those used for the regular-print, operational version of the test. Once the test administrator directions have been revised and approved by the CDE, they will be reviewed for use with the large-print edition and adjustments will be made, if necessary, in the page references and other directions. This version of the instructions will also be presented to the State for review and approval.

Test administrator instructions for the CD-ROM will, in addition to mirroring the standard administration manuals, include any instructions required for teachers to load and operate the CD-ROM. Again, this manual will be presented to the CDE for review and approval.

8. Component Task 8: Pre-Identification and Ordering (CST, CAPA, CMA, STS, NRT)

The STAR Management System allows district STAR coordinators to set up test administrations, order test materials, submit and update Pre-ID data, download early testing results and make demographic data corrections.

Organization System Access

A major user of this system is STAR TAC, whose staff provide assistance and information to district STAR coordinators on a daily basis. STAR TAC will collect updated designation and security forms. After receiving these forms for a district (or independent charter), STAR TAC activates the coordinator's User ID.

Simultaneously, STAR TAC is collecting and logging the updated designation and security forms from the districts.

When both forms are logged as received, the district STAR coordinator user access is activated by the application. An e-mail will be sent to the district STAR coordinator user with his or her username and temporary password. The text of the e-mail will be approved by the CDE.

Test coordinators for independent charters will have the same system access as a district STAR coordinator.

The requirement for a signed security agreement as well as the forced password change helps make the Management System site secure.

District and Administration Calendars

District STAR coordinator users will use the Test Administration Setup module to set up their multiple-choice and writing administration test dates. District STAR coordinators enter the instructional day calendars representing the district's schools; one calendar per administration. Non-instructional dates will be saved and updatable until the district STAR coordinator has approved the initial order for the administration.

When setting up a multiple-choice administration, if a user has indicated that the administration includes Grade 4 or 7 testing, the application will try to automatically create the appropriate writing administration. If the test date of the first writing administration is an instructional day, then first writing administration is created. If the date is a non-instructional day but the test date of second writing administration is an instructional day, then second writing administration is created. If both test dates are non-instructional days, then a writing administration is not created automatically. STAR TAC users will have the ability to create writing administrations without creating a multiple-choice administration. The wording of this user prompt will be:

> **Will you be administering Grade 4 or 7 tests to any students on this administration?**

8. A. Pre-Identification Process

8. A1. CDS Master File

As described above, the CDE will deliver a baseline version of the CDS Master File in early July prior to the administration year. Subsequent CDS Master File updates will be merged into the database during the administration year, so that CDS data used for Pre-ID edits is

consistent with the CDE file on record. Historically, the maintenance of the CDS data is complex. Different groups within the CDE are involved in the maintenance of the file and the data is dynamic, so deliveries of updates are staggered. Charter school status is the most dynamic aspect of the CDS Master File maintenance, but there are also cases each year where school districts merge or split. There are also schools added and deleted each year. At the same time, STAR TAC is collecting more recent and accurate information from the districts, and the Management System retains relevant data from the previous administration year. As a result, the merging of CDS updates with information in the management database requires a combination of programmatic and manual synchronization efforts.

Only district STAR coordinators and STAR TAC representatives are permitted to submit and modify Pre-ID data in the STAR Management System. The user must select the district for which Pre-ID data management is being performed. The district STAR coordinators must accept the terms of Pre-ID service before they can manage their Pre-ID data.

8. A2. Procedures Used to Verify the Completeness of Demographic Information Submitted by LEAs

While not required, the district STAR coordinator can submit the Pre-ID file to CSIS first to identify and correct any deficiencies. ETS and CSIS will conduct sufficiently detailed parallel systems testing prior to deployment in order to confirm consistent edits. To accomplish this, ETS and CSIS will compare and synchronize edit specifications and also load the same test files into both systems to compare the results for consistency. ETS will accept files from districts but will not accept district files from CSIS directly.

8. A3. Procedures to be Used to Load the Demographic Information onto a Secure, Interactive, Internet-Accessible Database

The user will be able to submit Pre-ID data using a pre-formatted student file. The layout of the file, as well as the format and permissible values of each field are specified by the CDE. A district STAR coordinator will have the option of downloading the previous year's Pre-ID data from the Management System to use as a starting point. Alternatively, the district STAR coordinator can start with the district's database and extract the data in the required format.

The Management System accepts Pre-ID files in fixed ASCII, comma delimited or MS-Excel formats. For MS-Excel, ETS provides a template.

When the user elects to submit a Pre-ID file, the application will provide the user with a "browse" mechanism to locate the student file locally for submission.

Once the file is submitted, the Management system will do the following to process the file:

- Determine the file's record format and database mapping rules; If the file does not pass pre-validation edits, the user will be informed to make corrections and resubmit.
- Perform validation edits for all fields in each student record in the file; Invalid records will be stored in the resolution table until corrected by the user.

A submitted file can be deleted so that a new file can be submitted. ETS recommends that a district STAR coordinator make global corrections to a file offline and then resubmit the file. However, ETS will support global correction capabilities in the Management System if required.

8. A4. Validation Process that Assures Valid and Complete Codes

Pre-ID Edit Rules

The Pre-ID edit specifications document will be approved annually by the CDE.

ETS's STAR Pre-ID process will support a multi-level approach to edit checks and error resolution in order to communicate errors to the district STAR coordinators while maximizing the ability to process and print Pre-ID data.

There are three types of fields: critical, intermediate and minor. The CDE-approved edit specifications will document the field type for each field.

County-District-School (CDS) codes and names will be checked for consistency with the CDS Master File. Unrecognized CDS codes will be rejected.

The following rules will apply:

- All failed edits will be summarized in the error summary log. The summary of errors will be organized by the type of field: critical, intermediate and minor. There will also be a detailed report itemizing errors by record number and field.
- For a given record, if any edit fails on one or more critical or intermediate fields, the records will be put in the resolution table instead of in the normal database tables. The resolution table stores the data as uploaded.
- For a given record, if all edits pass for all critical and intermediate fields, the records will be accepted and stored in the normal database tables.
- For records in the resolution table, the user has the ability to edit them online. When the record is selected for editing, all failed edits will be noted on the screen. Invalid values that the edit page could not display are converted to blanks and highlighted.
- For the file submission as a whole, the district STAR coordinator will also receive edit results relating to file-level thresholds for errors/omissions for certain key fields.
- For records still in the resolution table at the time of the Pre-ID cutoff:
    - A batch job will move them into the normal database table if only intermediate and minor edits have failed. Invalid characters in those intermediate and minor fields will be converted to blank values.
    - If, however, a record in the resolution table has a failed edit on a critical field, the record will remain in the resolution table and will not move to the normal database tables.
    - Once that batch job completes, another batch job pulls all the accepted records from the normal database tables whose Pre-ID cutoff time has passed and extracts them for document or label printing.
- ETS will conduct a joint validation test with CSIS to assure rule and data format consistency.

With this strategy, Pre-ID records with tolerable errors can still be processed. Districts have opportunities later in the test administration cycle to correct tolerable demographic data.

8. A5. Procedures to be Used to Notify LEAs that Data are Incomplete or Include Invalid Entries

The system will accept the file and communicate to the user the need to return to the site later (well within the required two working days) to view the results of file processing. If the district has provided the system their e-mail address, ETS will also send an e-mail to inform the district that file processing is complete. If not, the district will be alerted that they do not have an e-mail address to communicate to them. This will not stop the processing of the file.

STAR TAC will assist district STAR coordinators with their Pre-ID file submissions on an as-needed basis. STAR TAC staff will be trained to coach users through the process and have

the ability to log in on behalf of the district STAR coordinators as needed. A Pre-ID User Guide is also approved by the CDE and published on startest.org to guide users through the process.

8. A6. Procedures to be Used to Supply LEAs with Pre-Identification Labels

Districts will have the option to pre-identify students for each of the test administration types (CST, NRT, Writing, CAPA, CMA, and STS), using pre-identified documents or labels. The following is a description of the process that will be used.

1. Districts will submit enrollment and Pre-ID files to ETS by administration and test type (CST/NRT, Writing, CAPA, CMA, and STS). The enrollment order will indicate whether the district chooses Pre-ID documents, labels, or no Pre-ID.
2. ETS will edit the files using approved edit criteria, and will compare the enrollment order counts to the Pre-ID counts.
3. Clean files will be transmitted to PEM via FTP. Pre-ID files and enrollment files for pre-identified documents are due to PEM 42 calendar days prior to the first test date. Pre-ID and enrollment files for Pre-ID labels are due to PEM 21 calendar days prior to the first test date. Districts also have the option of submitting a late Pre-ID file for labels.
4. PEM will edit the files' critical fields for valid values, and will work with STAR TAC to resolve any Pre-ID or enrollment edits. If necessary, the files will be updated and re-transmitted.
5. Clean enrollment orders will be loaded into the order system where distribution rules will be used to calculate material quantities and overages.
6. Depending on the method of pre-identification selected by the district, the Pre-ID file will be used to produce adhesive barcode labels or to pre-print student demographic information on answer documents.
7. After printing, Pre-ID adhesive labels will be assembled alphabetically by grade and delivery name on the Pre-ID file and sent as a separate shipment from other testing materials to the district. Districts will receive Pre-ID labels within five to ten working days of ETS having received an accurate pre-identification file from the district. The district or school will be responsible for applying the Pre-ID label and matching the correct student to the label.
8. For those districts choosing pre-identified answer documents, after printing and packaging, their documents will be sent with the rest of their testing materials. The pre-identified documents will be sorted alphabetically by grade and delivery name on the Pre-ID file.
9. Information not provided in a Pre-ID file will be hand-coded at the time of testing. Full demographics will reside on the CST multiple-choice grade-level answer documents, CAPA, STS, and CMA answer documents. Only those demographic categories required for matching will reside on the CST Writing answer document, end-of-course answer documents and the NRT answer documents, thereby reducing the burden of gridding on the districts.

In the case of students pre-identified for all demographic categories, there will be no need for the student or school personnel to grid those fields on the test document. However, any field not included in the Pre-ID file may be gridded during testing.

The Pre-ID data will be applied to the student record after scanning, and will aid in the matching of student documents at those grade levels where students are assessed on more than one document (such as, Grade 4 CST multiple choice and Grade 4 Writing).

8. A7. Capacity to Modify Pre-Identification Files and Student Data Files

District STAR coordinators also have the ability to insert, update and delete Pre-ID records interactively in Management System Web pages. Updates and deletions can be made to records entered online or via the batch upload process.

The online edit errors are presented to the user interactively on the edit resolution page. Fields in error are highlighted for easy identification.

8. A8. Non-Tested Student Demographic Information

ETS will collect non-tester demographics at the time of testing through hand-coding. However, instead of using complete answer documents, ETS will design and produce a demographic-only intake sheet. This two-sided, scannable single sheet will include all of the same demographic categories that are on the answer documents for that administration. Categories and fields will be formatted to be comparable to the answer documents, making it easy for the teachers and district STAR coordinators to complete the forms.

To accomplish this, between 61,000 and 63,000 copies of the demographic intake sheet each year, including overage will be printed. Distribution quantities will be based on district enrollment and the previous year's State average for percent non-testing. Districts will receive the sheets with their other testing materials and will return the sheets for processing after testing with their scorable materials. PEM will scan the demographic intake sheets along with the district's answer sheets. Data captured on the demographic intake sheets will be provided along with assessed student data.

In the event that districts are able to identify non-testers in advance of the testing window, they may do so by adding those students to their Pre-ID files. The Pre-ID file layout would need to be modified to include such a designation, but would save teachers and district STAR coordinators from hand-coding demographic information during testing.

8. B. Ordering

8. B1. Process for Verifying that a Fully Executed Security Agreement has been Received Annually from each District STAR Coordinator Prior to Ordering

As described above in Section 8, the ordering process does not start until a fully-executed security agreement and designation form have been received.

Once the test administrations have been set up and the corresponding dates have been calculated, the district STAR coordinator is ready to order test booklets, answer documents and ancillary materials for the test administration. The ordering capability is available to districts each October, in conjunction with the CDE-approved ordering user manual and ordering workshops.

Before entering orders, the district STAR coordinator will confirm or update the shipping address and enter his or her working day calendar. Updates to the calendar will be allowed until the district STAR coordinator has approved an order for any administration.

This working day calendar will be applied to all test administrations for the district or independent charter. Delivery and return dates will then be calculated, as well as cutoff dates.

8. B2. Process for Obtaining Orders for all Testing Materials Needed for the Administration of the Tests

The district STAR coordinator will select the administration for which to submit orders and then select the schools for which to submit orders. The system will then present a Web page for the district STAR coordinator to enter enrollment quantities by grade and course for each school. After submitting the quantities, the district STAR coordinator will be required to approve the

order. After district STAR coordinator approval of the order, STAR TAC will approve the order. Then, at the appropriate time to meet the calculated delivery start date, the order will be fulfilled and shipped to the district (or independent charter). The fulfilled order shipment will include 10 percent more materials than ordered at the school level, and another five percent more material (than the sum of orders across all schools in the district) sent to the district. These overages will not applied to braille and large-print materials.

8. B3. Process for Tracking and Logging Orders

STAR TAC will monitor the ordering status of each district using a dashboard capability in the Management System. STAR TAC will focus on districts with orders that need attention. STAR TAC makes every effort to get all orders approved by the State-regulated deadline in December, but continues to work with districts to get orders submitted after the deadline.

District STAR coordinators can make requests for supplemental orders through the TAC. ETS will make sure that districts have everything that they need on testing day. While there are cutoff dates for ordering and Pre-IDs to encourage timely response on the part of the districts, STAR TAC will be very flexible in supporting last-minute requests.

During the approval process, STAR TAC staff will compare current order quantities against the previous year's order quantity to assure reasonableness of the order. The quantities ordered from the previous year will be summarized by district, school and grade in a spreadsheet.

8. B4. Process for Billing LEAs for Excessive Orders of Materials

In order to encourage districts to order only the test materials that they need, ETS will implement an excessive order billing policy as allowed by the contract. The algorithm will calculate the total cost of materials ordered by the district based on the prices and quantities of each product. Ninety percent of that value is the minimum cost of materials that the district must use in order to avoid the excessive ordering charge. The algorithm will then calculate the total cost of materials used by the district. If the minimum cost of materials to be used exceeds those actually used, but more than $100, the district is charged the difference.

9. Component Task 9: Test Materials Production and Packaging (CST, CAPA, CMA, STS, NRT)

9. A. Test Materials Production

[text deleted for Scope of Work]

9. B. CDE Copies

The CDE has final responsibility for the quality of STAR materials. ETS will assure that all the materials detailed below will be produced in the proper quantities and formats, reproduced on CD-ROM and delivered to the CDE within the prescribed timeframe (within four weeks of completion of materials production). Copies for the CDE will be produced as part of a normal quality-monitored production run, packaged and delivered via an approved, secure and traceable shipper to the designated CDE recipients. ETS will deliver to the CDE:

Two sets of all versions of test booklets, including:

- Booklets with built-in reference sheets
- CAPA Examiner's Manuals
- Two copies of all answer documents
- Two sets of rulers
- Two sets of CAPA stimulus cards
- One copy of all braille test booklets
- Ten copies of each DFA
- Twenty-five copies of all coordinator manuals

Additional copies of all of the listed materials will be available for the CDE's needs and will be provided to the CDE at their request.

9. C. Ancillary Test Materials

To increase the efficiency of communications between the LEAs and schools during the test administration, and to encourage that tests be administered in a consistent manner, ETS will annually review with the CDE all STAR Program documentation and update the materials based on the needs of the CDE and district STAR coordinators, academic standards, and other requirements and criteria.

All documents will adhere to the style and usage standards of the APA Publication Manual and the CDE Style Manual and refer to the ETS/STAR Terminology List when necessary, so that materials sent to LEAs use the same terminology and language as that used for STAR tests.

ETS will continue to produce hardcopy versions of test administration support materials for:

- *Directions for Administration* (*DFA*) for Grades 2 through 11
- The *District and Test Site Coordinator Manual* (for Pre-Test Workshop)
- *Post-Test Manual*
- Guides for the STAR Management System
- CAPA, CMA, and STS *Examiner's Manuals*

ETS will provide documentation for the CAPA, CMA, and STS, and the corresponding examiner's manual.

Directions for Administration

ETS will produce one *DFA* for most grades as well as one manual for high school–level testing:

Grade 2	Grade 6
Grade 3	Grade 7
Grade 4	Grade 8
Grade 5	Grades 9–11

*DFA*s include instructions for completing demographic information and placing linking labels. They also include information for the security and the return of test materials.

District and Test Site Coordinator Manual

The *District and Test Site Coordinator Manual* instructs district and test site coordinators in the administration of the STAR tests. The manual contains procedures and forms that are specific to the coordinators' roles in the STAR Program and serves as a complete reference guide.

Post-Test Guide Technical Information

The current *Post-Test Guide Technical Information for District and Site STAR Test Coordinators and Research Specialists* supplements the post-test workshops. It will be written to enhance understanding of the reports that are generated from scoring and analyzing STAR results on the CST, NRT, and CAPA tests, to instruct in the derivation and use of the California Reading List Number, and to offer procedures for data correction.

STAR Management User Guides

ETS will also produce documentation that shows district STAR coordinators procedures for using the STAR Management System, an outline system that allows districts to order materials, submit Pre-ID files, make demographic edits, such as how to order, how to make data corrections. In particular, the user guides will show the coordinators how to set up test administrations, and how to submit Pre-ID files. These user guides will be available as PDF files only and can be downloaded from startest.org.

CAPA Examiner's Manual

ETS will produce a manual for the administration of the CAPA exams, the *CAPA Examiner's Manual*. One manual is used whether CAPA Level I, Level II, Level III, Level IV, or Level V is administered.

The *CAPA Examiner's Manual* describes the CAPA Program, qualifications for administration, determination of CAPA levels, and instructions for examiners' scoring of the test. It also describes:

- Recording responses
- Using stimulus cards for certain test questions
- Using manipulative lists for certain test questions
- Guidelines for administration
- Adapting tasks

ETS psychometricians will produce the stimulus cards and manipulative lists needed to administer CAPA.

Minimum Quantities of Paper Products

For efficiency and consistency, ETS will produce standard manuals and instructions as all-inclusive documents to support both standard test versions and special versions (large-print, braille and CD-ROM). The ancillary materials will have clearly delineated sections to support the unique instructions for delivering these assessments.

Table 27. Proposed Minimum Quantities of Ancillary Paper Materials

Material	To Each LEA	To Each School
*DFA*s	2	1 per 20 tests
District and Test Site Coordinator Manuals	1	1
Post-Test Guides	1	
CAPA Examiner's Manuals	0	By request

Directions for Administration

There will be one copy of the instructions per 20 test takers. A 10 percent overage will be added to all school orders, as well as a five percent overage for districts, and an additional amount to be held in reserve to support supplemental requests. In addition, the posting of electronic versions should support any extraordinary needs.

District and Test Site Coordinators Manuals

ETS will provide two per district and one per school, with some overage to support additional requests by administrative personnel. The electronic versions posted on the STAR Web site are print-ready and capable of supporting any extraordinary demands.

CAPA Examiner's Manuals

The *CAPA Examiner's Manual* will be produced at a much lower ratio than 20:1, reflecting the population dispersion of candidates for this assessment. Since the *CAPA Examiner's Manual* contains test questions, district STAR coordinators will order one for each examiner.

Ancillary Materials Production Specifics

All ancillary materials will be printed, and converted into PDF and HTML files. *DFA*s, coordinators manuals, and the CAPA manual will be posted to startest.org by February 1 of each year; hardcopy versions will be shipped with testing materials. Only non-secure materials or materials that have been edited to remove secure sections are posted. Posted materials, such as *DFA*s and the non-secure *CAPA Examiner's Manual,* will be accessible using Web Braille.

9. D Packaging

9. D1. Quality Control Procedures

Printing Quality Control Procedures

An ETS Printing Quality Control Specialist will be on site through all stages of production to assure the quality of all products.

The general process requires all ETS print vendors to perform a quality check on all materials produced at all stages of print manufacturing. The quality checks are performed at the Prepress, Press, Bindery, and Packaging/Shipping stages. ETS will use a required Quality Control Check List to assure the vendors' adherence to quality procedures.

Scorable Materials Production

Sections 7. B and C include additional information about the document design and construction process.

PEM will produce scannable documents using standardized processes set by ISO (International Organization for Standardization), that promote efficiency and improved quality. Adherence to the process standards set by ISO means that forms and printing equipment are kept to a high level of precision.

Prior to shipment from the printing facility, the Quality Department will approve all samples. Any form that does not meet the scanner reflectivity tolerance guidelines is rejected. Forms that fail to meet these quality control criteria are reprinted.

9. D2. Test Materials must be Packaged, Labeled and Shipped

Packaging

PEM will package test booklets, scorable documents, and DFAs. ETS will distribute coordinator manuals at workshops and ship the manuals to districts that do not send representatives to workshops.

With the exception of Pre-ID labels, all test materials for a test administration window will be shipped at one time. All materials will be packaged for each school and shipped to the district. Pre-ID labels will be sent as separate shipments to allow districts more time to submit pre-identifying information.

Because many districts have multi-track calendars that require testing in more than one test administration window, it will be necessary to make more than one shipment to some districts.

The materials list is a requirements document that specifies anticipated page counts, order quantities, distribution quantities, and processing quantities for each item type by year.

Distribution rules will be used to calculate material quantities and overages, to provide 10 percent overage for every school testing and give five percent overage for every district, based on the districts' total order for each grade.

Boxes will be packaged by assessment and grade for each test site, and sent to the district. For example, CST/NRT materials are boxed separately by grade from the CAPA test materials. As STS and CMA test materials are required, they will be packaged separately from other assessment materials by grade. The contents of each shipment will be clearly labeled.

Box 1 of each district- or county-office shipment will include:

- Return freight kits for scorable and non-scorable materials
- Directions for inventorying the materials and for notifying STAR TAC of any missing materials or shortages
- A set of packing lists for all school shipments within the district or county office
- A packing list for the district or county overage materials listed in the order in which they are packed
- A pallet detail report for those shipments that include two or more pallets

Box 1 of each school shipment will include:

- Return freight kits for scorable and non-scorable materials
- A packing list with materials listed in the order in which they are packed

The boxes will be labeled numerically to correspond with the packing list (Box 1 of 20, Box 2, Box 3, Box 4...Box 20 of 20), so that materials for a particular assessment and grade level can be identified upon receipt at the test site.

Box Specifications

PEM will use specifications for box construction so that the boxes used for STAR are extremely sturdy and durable. PEM will use double-walled, reusable boxes to both withstand the rigors of handling by the carriers during distribution to districts, and to protect the test materials when they are shipped back to PEM for processing.

California school districts receive many shipments of materials during the school year. Although the boxes are labeled with program information, it is critical that the district STAR coordinator be able to locate Box 1 of each shipment as soon as possible. Box 1 contains the packing list and other important information needed to facilitate handling.

Using suggestions from districts, PEM will institute a procedure beginning in 2007 to package all materials for Box 1 in a white, double-walled box. The white box will be easily recognized in the shipment as Box 1.

Special Services to Districts

Where possible, PEM provides:

- Accommodations for special needs and space
- Pallet jack or other equipment necessary for districts without a dock or proper equipment
- Alternate carrier arrangements so that testing materials reach the more remote areas of the State on time

PEM will use the special comments section on the enrollment order to capture requests for proper delivery (no dock, need assistance, etc.). If there are any questions about the comments entered on the enrollment order, STAR TAC will follow up with the district STAR coordinator before test materials are shipped for that district.

Additional Orders

When districts need additional materials, requests for additional materials will be processed as long as shipments to other districts are not delayed. Having the additional orders fulfilled using the main packaging and distribution system will allow PEM to consistently and effectively respond to requests for additional materials.

Errors and shortages in orders will be filled and shipped to districts within two business days of notification.

Packaging and Distribution System

PEM will utilize the state-of-the-art Oracle Packaging and Distribution system. Oracle's Packaging and Distribution system uses barcode-identified packaging components. Barcodes will identify item type, boxes, orders, pallets, and shipments.

Pre-Packaging Process

Once the test material specifications are finalized with the CDE, ETS will order ancillary testing materials, such as math rulers and reference sheets, special versions, and nonscannable test booklets from its certified vendors. PEM will complete internal purchase orders for the scannable documents, and printing will begin according to the project schedule.

All test booklets for the STAR Program will have a barcode printed on the back cover of the booklet which identifies the booklet type (grade level and subject matter, if applicable). Each

booklet type will be spiraled and shrink-wrapped in packages of fives and twenties prior to moving the materials to the packaging line. The Grade 2 CST versions are not spiraled. Versions are assigned by school since the questions are read aloud to the students.

Upon receipt of the printed materials, PEM will perform the following steps:

1. Count materials to confirm receipt of all items ordered.
2. Enter quantities into the Oracle Packaging and Distribution system.
3. Transfer materials to the production floor for packaging.

Prior to moving a project into production, the following steps will occur:

1. A project specification form will be created. The form will contain all information required for Oracle Packaging and Distribution. The form will reside in the project team's central repository.
2. A packaging and distribution schedule will be built to verify that all testing material will be received in districts according to the project requirements.
3. For each site, a transportation file will be created, which consists of requested quantities of each material type along with calculated overages.

Project-specific system tests will be conducted to verify that the system is functioning accurately. The test routines will run end-to-end from the order-entry stage through the final packaged product.

Customized items such as Pre-ID answer documents, Pre-ID labels, inbound and return labels, school and grade ID sheets, school master file sheets, and Pre-ID rosters will be created during the pre-packaging process.

10. Component Task 10: Delivery and Collection of Test Materials (CST, CAPA, CMA, STS, NRT)

10. A. Delivery of Test Materials

As described in Section 9. D, PEM will utilize the state-of-the-art Oracle Packaging and Distribution system to manage inventory control and maintain accuracy in the packaging and distribution of the STAR Program materials.

Using this system, PEM will provide districts with a set of shipping documents that includes:

- A packing list for the district overage materials
- A packing list for each school's materials
- A pallet map that shows the identity and pallet assignment of each carton

If a district or school should misplace any of the documentation during inventory check-in, PEM will send replacement documentation via e-mail to prevent delays in distribution of materials.

10. A1. Inventory Control Procedures

Using the Oracle Packaging and Distribution system, school district orders will be entered into a database, and order information will be formatted for the packaging processes by a proprietary order-entry program. As each order is posted on a barcode-scanner screen, packaging staff will scan barcoded items and assign the materials to a unique carton.

10. A2. Handling Shipments for LEAs that have Multiple Test Administration Periods

Because many districts have multi-track calendars that require testing in more than one testing window, it is necessary to make more than one shipment to some districts. ETS and PEM will work with the districts to deliver test materials to the district in response to their testing dates.

10. A4. Assigning Unique Identifiers to Every Test Booklet

All test booklets for the STAR Program will have a barcode printed on the back cover which identifies the booklet type (grade level and subject matter if applicable). Each booklet type will be shrink-wrapped in packages of fives and twenties prior to moving the materials to the packaging line. Each package will have a barcode label applied to it, which will be used to pack and log the number of documents of each type sent to each district, by test site. As secure materials are returned from each district PEM will verify that all test booklets were returned. Section 10. B3 describes procedures for notifying LEAs of discrepancies between the quantities of secure materials that were originally shipped and the quantity of secure materials received by PEM for processing.

10. A5. Procedures and Tracking Processes

PEM will meet the CDE requirements for multiple-choice materials to arrive in the districts 10 to 20 working days before the first testing day of each administration, and writing test materials to arrive in the districts 5 to 10 working days before the writing test administration.

10. B. Collection of Test Materials

10. B1. Procedures for Picking Up All Scorable and Non-Scorable Secure Materials from LEAs

In order to expedite the return process, PEM will assign each district a carrier to contact for pickup when materials are ready for return shipment. Because the STAR Program

Management Team knows each district's carrier assignment, ETS has access to each district's return-shipment tracking numbers. The tracking number allows the monitoring of return-shipment activity. PEM will maintain contact with the carriers to address emergencies or other situations, such as bad weather or district-specific needs.

Scorable and non-scorable materials must be returned within five working days after the last day for each test administration period. For the CST Writing test materials, PEM expects return within no more than two working days after the makeup day for each administration. PEM will closely monitor the return of materials and will notify the STAR Technical Assistance Center of any districts that have not returned their materials. STAR TAC will contact the district STAR coordinators and work with them to facilitate the return of the test materials. ETS will collaborate with the County Offices of Education to work onsite with districts to return materials in a timely manner.

In the packaging process, PEM will include freight return kits for scorable and non-scorable materials for use by the District and Test Site Coordinators. The freight return kits will contain color-coded labels identifying scorable and non-scorable materials. The label will also contain barcoded information identifying the school and district. When test site coordinators pack their materials for return to the district, they are required to apply the appropriate labels and number the cartons (such as 1 of 2, 2 of 2). Upon receipt of the materials in the district, the district STAR coordinators are required to complete the "total shipment from this district" information on the label.

The use of the color-coded labels streamlines the return process at PEM. All scorable materials will be delivered to their scanning and scoring facilities in Iowa City, IA. The non-scorable materials, including test booklets, are returned to the Security Processing Department in PEM's Cedar Rapids, IA facility.

10. B2. Processing of Returned Materials

Upon receipt of the test materials, PEM will utilize a precise inventory and test processing system in addition to quality assurance procedures to maintain an up-to-date accounting of all the testing materials within their facilities.

As PEM receives test materials, they remove the materials from the shipping cartons and carefully examine each shipment for a number of conditions, including physical damage, shipping errors, and omissions.

As PEM batches materials for scanning, they also do a visual inspection to compare the number of students recorded on the School and Grade Identification Sheet (SGID or Header Sheets) to the number of answer documents in the stack.

PEM's image scanning process provides the ability to capture security information electronically and to do the following:

- Compare scorable material quantities reported on header sheets to actual documents scored.
- Follow up on any missing shipments or quantities appearing to be less than expected with a phone call by the PEM Program Management Team to the district. STAR TAC staff will contact the district for further resolution.

All secure materials will be checked into the PEM Cedar Rapids facility by scanning the barcode label on each of the returned cartons. The materials in each box will be counted and returned to the original box for storage. The quantity of test booklets received by PEM, including the scanned counts of Grades 2 and 3 scorable documents, will be compared to the quantity that was assigned and sent to each district and school.

10. B3. Notifying LEAs of Discrepancies Between the Quantities of Secure Materials

PEM will send reports detailing secure materials received back from the districts or schools to STAR TAC. Follow up with the districts is handled by STAR TAC. PEM will provide the CDE with an electronic file showing the final resolutions of discrepancies no later than September 20 of each year. The format of the file used for the 2007 through 2009 administrations will be similar to the file format currently used.

10. B4. Procedures for the Secure Destruction of Secure Materials

After secure materials (including test booklets and examiner's manuals) are processed, they will be returned to their original boxes for storage, and palletized and placed in PEM's secure warehouse facilities in Cedar Rapids, IA. Once all resolution is complete, PEM will request approval from ETS and the CDE to salvage the materials. PEM understands the importance of security, including during the salvaging process.

11. Component Task 11: Test Processing, Scoring, and Analysis (CST, CAPA, CMA, STS, NRT)

11. A. Quality Control and Assurance

11. A1. Quality Control Checks

Before any STAR documents are scanned, PEM will conduct a complete check of the scanning system. PEM's Software Quality Specialists (SQS) will create a test deck for every test and form. Each test deck will consist of 25 answer documents gridded to cover response ranges, demographic data, blanks, double grids, and other responses. Mock students will be created to verify that each gridding possibility is processed correctly by the scanning program. The output file created will be thoroughly checked against each answer document after each stage to verify that the scanner is capturing marks correctly. When the program output is confirmed to match the expected results, a formal sign-off process will take place. A scan program release form will be signed and the scan program will be placed in the production environment under configuration management.

Quality Control of Image Editing

Prior to submitting any STAR operational documents through the image editing process, PEM will create a mock set of documents that will test all of the errors listed in the edit specifications. The set of test documents are used to verify that each image of the document is saved so that an editor will be able to review the documents though an interactive interface. The edits will be confirmed to show the appropriate error, the correct image to edit the item, and the appropriate problem and resolution text that will instruct the editor on the actions that should be taken.

Once the set of mock test documents is created, the following procedures are completed:

- Scan the set of test documents.
- Verify that the images from the documents are saved correctly.
- Verify the appropriate problem and resolution text displays for each type of error.
- The image edit system submits the post-edit program.
- If the post-edit identifies errors that still require correcting, make changes and resubmit the post-edit program.
- Print a listing of the post-edit file, the correction card file and the original scan file.
- Check correction cards against the post file for corrections made. The post file will have all keyed corrections and any defaults from the edit specifications.

In addition to the quality control checks carried out in Scanning and Image Editing, the following manual quality checks will be conducted to verify the answer documents are correctly attributed to the students, schools, districts, and subgroups:

- Building counts are compared to the District Master File Sheets.
- Document counts are compared to the School Master File Sheets.
- Document counts are compared to the School and Grade Identification Sheets.
- All districts/buildings are compared to the CDE CDS Master File.

If any discrepancies are identified in the steps outlined above, the PEM Product Line staff will follow up with the districts for resolution.

The STAR Management System will create expedited results files for districts and extracts for Web reporting and corresponding student data files from the scored data.

All production software programs associated with the STAR Management System are subject to the following quality assurance processes prior to production deployment:

- CDE-approved requirements and specifications
- Software configuration management strictures so that the software that is developed is the same as the software that is tested and ultimately deployed
- Software testing by a development group and then by a separate testing group, including regression and performance testing
- Defect tracking so that defects can be resolved efficiently
- Operational Readiness Reviews, where ETS Scoring, Reporting and Technology stakeholders assess deployment readiness of software releases prior to deployment approval

Aggregated results to be published on any paper or online reports will also checked by the Data Quality Services department of Scoring, Reporting and Technology in conjunction with the Research Division's quality control process.

Prior to processing operational answer sheets and executing subsequent data processing programs, ETS will conduct an end-to-end test. ETS will prepare approximately 700 test cases covering all tests and many scenarios designed to exercise particular business rule logic. ETS will grid answer sheets for those 700 test cases. They are then scanned, scored and aggregated. The results at various inspection points will be checked by Research and Data Quality Services staff. Additionally, a post-scan test file of approximately 50,000 records will be scored and aggregated to test a broader range of scoring and aggregation scenarios. These procedures assure that students and districts get the correct scores.

11. A2. Handling of Answer Documents

All secure materials will be checked into the Cedar Rapids facility by scanning the barcode label on each of the returned cartons. The materials in each box will be counted by scanning the barcode that identifies the material type on each of the documents. After the contents of the box are scanned, they will be returned to the original box for storage. The quantity of test booklets, including the scanned counts of Grades 2 and 3 scorable documents received at PEM will be compared to the quantity that was assigned and sent to each district and school. PEM sends reports detailing materials received back from the districts/schools to STAR TAC. Followup with the districts is handled by STAR TAC. PEM will provide the CDE with an electronic file showing the final resolutions of discrepancies no later than September 20 of each year.

11. B. Test Processing

11. B1. Timeline for Test Processing, Identify Personnel, and any Subcontractors Involved in the Process

ETS Program Managers will be responsible for monitoring all test processing activities assigned to PEM and work with PEM Test Administration Coordinator and Subcontractor Project Coordinator for STAR to verify that all PEM tasks are completed on time and according to requirements. Both will rely heavily on other personnel in their organizations to complete test processing successfully.

District STAR coordinators will return all multiple-choice scorable and nonscorable materials to the STAR Scoring and Processing Centers no more than five working days after the completion of testing for each administration. District STAR coordinators will return all Grade 4 and 7 scorable and nonscorable materials for the Writing application to the STAR Scoring and Processing Centers no more than one working day after the make-up date for the Writing test administration. All test materials will be returned via ground delivery to PEM.

Once the scorable materials have been scanned, edited, and scored, and have cleared the clean post process, the results will be submitted to ETS to post on the secure STAR Web site. For districts using the pre-equated test form, electronic results will be available within three weeks after the scorable materials are received at the STAR Scoring and Processing Center. For districts using the post-equated test form, the electronic results will be available within three to six weeks after the scorable materials are received at the STAR Scoring and Processing Center. Refer to the Schedule of Deliverables and Activities in Section 3.1 A for a detailed timeline for test processing.

11. B2. Editing All Answer Documents

Three opportunities for demographic data to be edited will be provided:

- After scanning, by PEM online editors
- After online editing, by district STAR coordinators (demographic edit)
- After paper reporting, by district STAR coordinators

The demographic edit process is described in Section 11. B3.

ETS will use error conditions to indicate:

- Any blank in the first three positions of the first or last name
- Multiple marks in any position of the first name, last name, or MI (middle initial)
- Part of the date of birth is blank or contains multiple marks
- Grade is blank or contains multiple marks
- Student ID contains multiple marks
- Test version contains multiple marks
- Blank document
- Lithocodes do not match

As part of the edit process, name, grade, birth date, and gender will be edited by PEM's online editors when multiple grids for a field have been filled in, or the data is invalid (e.g., date of birth is out of range). In these cases the editor is presented with the image clip of the data in question and makes a determination if the student truly did make multiple grids or perhaps hand-coded the name or date correctly, but gridded the incorrect corresponding ovals.

Districts will have the ability to update missing demographic data including name, grade, birth date, and gender, and other demographic data after the paper score reports are sent to the districts, as described in Section 11. B5.

11. B3. Demographic Edits

11. B3a. Procedure for Flagging and Notifying Schools and/or LEAs of Answer Documents Missing a Certain Percentage of Student Demographic Data

After the data has passed the specified edits and the file and the district or district testing in multiple test administrations has been combined, the file will be edited to check if missing required student demographic data meets or exceeds the STAR Program tolerance of three percent and the number of answer documents submitted for scoring is 10 or more.

The following required student demographic data will be edited:

- Primary Ethnicity
- English Proficiency
- School Mobility
- Counted in October CBEDS – School
- Counted in October CBEDS – District
- Primary Disability Code
- National School Lunch Program Participation
- SSID number

11. B3b. Procedure for Obtaining Missing Information, Entering Information Into the Student Record, and Maintaining a File of Edited Demographic Areas by School

If the data exceeds the tolerance requirements, PEM will contact the district STAR coordinator and provide them with a description of the demographic edit problem. The district STAR coordinator will have three business days to provide the missing data. Districts can update their demographic data directly in PEM's SchoolHouse™ System Web site or through a paper process.

Once the district is finished with the updates, the updated data is pulled from the SchoolHouse™ Web site and applied to the scoring database where all of the student data resides. When the administration and/or district is complete, the data is extracted for reporting purposes.

11. B3c. Process for Developing and Providing an Electronic File to CDE Every Two Weeks Beginning Mid-June

For purposes of withholding apportionment monies to pay the district costs incurred for editing files for missing demographic data, PEM will deliver an electronic file to the CDE every week beginning in early June. The file will include the number of students within each district or County Office of Education for which demographic edits were required. The weekly files will include data only for districts and counties for which processing has been completed. The report for districts testing in multiple test administration windows will be held and forwarded after processing the last administration for each district. The report for each district or county will include school-by-school numbers and a total count for the district or county office.

11. B3d. Provide and Maintain a Secure Internet Site that the CDE can Access to Determine the Status of Each LEA's Demographic Edit Checks

In addition to the weekly file, the CDE will also have access to PEM's SchoolHouse™ System to determine the status of each district's demographic edit checks.

As described in the previous requirement, Section 11. B3c, PEM will also provide a weekly file to the CDE. This file will provide the following:

- The capacity to track where each LEA is in the edit process (being checked, failed edit, passed edit, etc.)
- The number and percent of student answer documents that failed the edit check at each school
- A summary of the number of student answer documents that failed the edit check in each LEA
- The number and percent of schools in each LEA that failed the edit check
- State level totals including the number of student answer documents that failed the edit check, as well as the number of schools and LEAs

11. B4. Matching Each Grade 4 and Grade 7 Student's MC ELA CST Score with ELA CST in Writing Score

The ELA CST multiple-choice score for each Grade 4 and Grade 7 student will be matched with his or her ELA CST writing test score. The documents will be matched within building/grade.

The match criteria are:

- ID Number (SSID number if present; if not, will use the student ID number)
- First name, last name, date of birth, and gender
- The multiple choice demographic data will be considered the demographic data of the record.

Scores for student documents that cannot be matched based on these criteria, will be reported separately. In addition, districts will receive an unmatched report with their reporting package listing the Grade 4 and 7 students for which there was a multiple-choice score but no writing score, and Grade 4 and 7 students for which there was a writing score but no multiple-choice score.

11. B5. Matching Records for Students who have Incomplete Test Results

Answer documents will be designed to produce a single complete record for each student, which will include demographic data and all test scores for that student. The following matrix identifies the type of answer document(s) that will be utilized at each grade level.

For those students in Grades 2, 5, and 6, demographic information and responses will be captured on a single, multi-page document. Students testing in Grades 3, 4, 7, 8, 9, 10, and 11 will be assessed using multiple answer documents.

All STAR answer documents will contain uniquely numbered lithocodes that are both scannable and eye readable. The lithocodes will allow all pages of the document to be linked throughout processing, even after the documents have been slit into single sheets for scanning.

For those students testing on more than one answer document, lithocodes will link their demographic and responses within a document, while matching criteria will be used to create a single record for all of the student's documents. PEM will use the matching criteria listed below.

- Grade 4 CST and Grade 4 Writing
- Grade 7 CST/NRT and Grade 7 Writing
- First — SSID number

- Second — First name, last name, date of birth, gender
- Grade 3 CST and NRT
- First — Pre-ID Barcode number
- Second — SSID ID number
- Third — First name, last name, date of birth, gender
- Grade level (8, 9, 10, or 11) and end-of-course tests (Math, Science, and/or World History)
- First — Pre-ID Barcode number
- Second — SSID number
- Third — First name, last name, date of birth, gender

CAPA. Twenty percent of the CAPA tests are scored more than once (two answer documents received). The answer document indicates whether the test was scored by the Examiner or the Observer. The two will come in under the same School and Grade ID sheets and need to be matched.

- First — SSID number
- Second — First name, last name, date of birth, gender

Data on STS grade-level answer documents and STS math end-of-course documents will be matched using the same matching criteria used to match CST grade level and end-of-course answer documents.

With the addition of end-of-course answer sheets to the STAR Program, matching for students at the higher levels will need to take place on up to four answer documents (grade level, Math, Science, and World History).For non-matched records, all student data and results are reported, but as individual records. The Pre-ID record layout will be modified to include fields for the end-of-course tests, so that end-of-course answer documents can be appropriately pre-identified.

A separate STS end-of-course math answer document will also be provided for the 2008 STS field test and 2009 operational administration. ETS will match STS grade-level data with STS end-of-course math data to create student records.

Districts will have the ability to update additional missing demographic data including name, grade, birth date, and gender after the paper score reports are sent to the districts. PEM will produce an Unmatched Report that is mailed to the districts with the score reports. The districts can update the missing data identified in the reports by accessing the STAR Management System. The time period for making the updates will be determined by the CDE and ETS. At the end of that time period, ETS will send PEM a file containing only the changed records. PEM will process the changes, load the data to the scoring engine, rescore if necessary, and aggregate all affected districts prior to delivery of the final Internet file to ETS. In addition, PEM will produce a student data file and reports upon request from the districts.

11. B6. Electronically Capturing and Storing the Answer Documents

After the answer documents have been scanned, edited, scored, and have cleared the clean post process, they will be palletized and placed in the secure storage facilities at PEM. The materials will be stored until October 31 of each year, at which time PEM will request permission to salvage the materials. After receiving CDE approval, the materials will be salvaged in a secure manner.

Due to the volume and size of the STAR answer documents processed, it is more cost-effective to store the paper documents for the life of the contract than it is to image nearly five million answer documents per year. PEM will store answer documents in original paper form.

11. B7. Producing an Electronic Mark Discrimination Report

PEM will create a Mark Discrimination report that will be delivered to the CDE no later than September 15 each year. The final specifications of this report will be mutually agreed upon by the CDE, ETS and PEM.

The Mark Discrimination Report will include any group greater than or equal to 10 students for which the number of answers changed exceeds the statewide average for the same grade and test by two standard deviations.

The Mark Discrimination Report will include the following:

- Number of students with changes
- Number of students in the group
- Average number of changes per student
- Percent of all responses changed
- Percent of responses changed from right to wrong
- Percent of responses changed from wrong to right
- Number of items

The following filters will be applied to the initial Mark Discrimination Report:

- Group was identified for two or more tests
- More than 50 percent of students in the group had changes
- Sixty-five percent or more of all changes were from wrong to right or 90 percent or more of changes on any one test were from wrong to right when the average number of changes was greater than or equal to 2

PEM will use the filtered Mark Discrimination report to produce a list of test groups in alphabetical order within grades within schools, schools within districts, and districts within counties. PEM will also produce student lists showing the students' item responses for the tests identified. The answer documents for the filtered list of test groups will be pulled and shipped to the CDE as soon as the groups are identified. Documents for each identified testing group will be packaged together with the shipping carton(s) clearly labeled.

PEM will work closely with the CDE and ETS to define a procedure for identifying answer documents that are randomly or pattern marked. PEM will also collaborate with ETS and the CDE to define the guidelines for reporting individual results for students with pattern or random marked answer documents as well as addressing these results in school, district, county, and state summary reports.

11. C. Scoring and Quality Assurance for the CSTs, CAPA, CMA, STS, and NRT

11. C1. Scored Answer Documents

All district names, school names, and CDS codes will be verified and confirmed with the CDS Master File provided by CDE.

11. C2. Develop and Produce Scoring Protocols and Programs for all Items and Other Scoring Materials

ETS will write scoring procedures and specifications that will help assure an error-free method of processing and scoring test materials. These include:

- **General Reporting Specifications** — defines various terms (for example, number of students enrolled, number of students tested, number of students with valid scores) and documents how to differentiate answer documents when no test items are marked
- **Score Key and Score Conversions** — describes the process of scoring and converting scores
- **Aggregation Rules** — describes how and when a school's results aggregate to the district level and then to the State level
- **What If…** — describes unusual and/or irregular situations and conditions discovered on receipt of used test materials and provides the action(s) to be taken
- **Edits** — describes edits, defaults and solutions to errors encountered during the data capture stage
- **Reporting Cluster Names and Item Numbers** — describes the names of the reporting clusters for each section of the test and which items make up that cluster

These procedures and specifications will be reviewed by the CDE; when both parties are satisfied that they are correct, the CDE will issue formal approvals. ETS will then use the approved procedures and specifications to program the scoring system.

11. C3. Process to Provide all Scoring Specifications

The previous year's scoring specifications will be used as a baseline for next year's scoring specifications. To ascertain changes made to these specifications, ETS will review all pertinent documentation including changes to the answer documents. Weekly (or more often) meetings will be held with CDE staff to discuss the revised specifications. During these interactions, a working, updated template of the specifications will submitted electronically to the CDE. When both parties are in agreement to the finality of the specification, the CDE will issue a formal approval of the scoring specifications.

Scoring keys will be verified through at two locations. PEM will verify their scoring internally. ETS will independently verify its scoring of the data and then compares the two results. Any discrepancies will then be resolved. The entire scoring system will initially be tested using a test deck that systematically varies expected cases and cases that occur rarely in real data. Following this, Classical Item analyses will be run on an early sample of data to further verify the scoring of actual data to provide an additional check of the keys. Following the equating results longitudinal data from complete districts will analyzed for reasonableness of results for all tests. This analysis is repeated with the “P1” file (90 percent of the data) to look at state tends and trends for the largest districts. These results will be provided to the CDE and jointly discussed. Any anomalies in the data will be investigated further and again jointly discussed days later. When satisfactory explanations for the results are obtained and both the CDE and ETS are comfortable with the results, the scores will be released.

11. C4. Process for Excluding Student Scores from Summary Reports

ETS will provide specifications to the CDE that document when to exclude student scores from summary reports. These specifications will include the logic for handling answer documents that, for example, indicate the student tested but marked no answers and was absent, not tested due to parent/guardian request, or did not complete the test due to illness. The methods

for handling other anomalies such as a Grade 8 math test where the specific math test is unknown, will also be covered in the specifications. The CDE will review the specifications and when both parties are satisfied that the specifications are accurate, the CDE will issue formal approvals. ETS will then use these approved specifications to program the reporting system.

11. C5. Scanning/Scoring Programs to Capture Codes to Indicate Changed Answers

PEM's scanning and scoring programs will capture and include codes in the electronic student records indicating answers that were changed from wrong to right or right to wrong. PEM will continue to provide this information for the 2007–2009 STAR Program.

11. C6. 4-Point Holistic Scoring Rubrics

PEM will use regional scoring with a distributed scoring subset to include California educators.

STAR Writing

PEM's Performance Scoring will score all Grade 4 and 7 Writing prompts.

Each student will produce a response to one writing prompt that will be scored as follows:

1. A 4-point holistic rubric pre-approved by the SBE will be used.
2. One professional reader who has passed PEM's rigorous scorer training will read and score each paper.
3. A minimum of 10 percent of the papers will be read by a second PEM trained professional reader to check the accuracy and consistency of scoring.

PEM's Performance Scoring will work with the CDE to determine how to mark unscorable documents.

Performance Scoring Regional Scoring Sites

PEM will use its network of scoring sites to provide image-based scoring services.

These sites will employ the same image-based scoring system used by California teachers who are scoring via PEM's Distributed method, and includes online training, allowing PEM's Performance Scoring to maintain the advantages of the Distributed scoring system while employing the resources of regional sites in a Regional Scoring model.

Along with the site-based scoring pool, PEM will utilize approximately 15 to 20 percent of the scorer population as distributed scorers. PEM proposes the inclusion of California educators as distributed scorers. The distributed scorers received the same level and quality of support as did scorers located at the regional sites. A fully staffed scoring support center will provide direction to the distributed scorers in terms of technical, content and human-resources-related issues. Along with the support center, a secure Web site is maintained and provides information and updates concerning the progress of the project.

Rangefinding Process

Each student response photocopied for review by the committees will be assigned a unique number for rangefinding purposes, and a corresponding log will be used to record important comments and decisions. Committees will systematically review the photocopied responses, prompt by prompt, determining and recording consensus scores, and making recommendations for the possible placement of papers within training sets.

Performance Scoring team members will keep a formal log of all papers discussed, recording all scores assigned along with any recommendations for the placement of papers in training sets. The log will be presented to the CDE after the rangefinding sessions for confirmation that the committee decisions and official scores were accurately recorded. The CDE and PEM representatives will sign the log to certify that the scores have been accurately recorded.

Rangefinding and Development of Training Materials

PEM's Performance Scoring will assume responsibility for the development of training materials, and will oversee and manage training and scoring processes for all STAR Writing prompts in Grades 4 and 7. PEM will use a set of procedures for developing training materials that involves the CDE, SBE test liaison, ARP representatives, and California teachers, and will result in comprehensive training materials so that PEM scorers will become deeply familiar and prepared for each prompt.

The first step in the development of training papers will be rangefinding committee meetings. ETS will work with the CDE and ARPs to develop a list of active California teachers at the grade levels being assessed and from various regions of the state. The committees will set the standards for interpreting the rubrics by determining consensus scores on a large sample of student responses.

PEM scoring staff will use these responses and corresponding consensus scores to construct comprehensive training courses for each prompt. The training course for each prompt will include a set of anchor papers to show a range of papers at each score point, two to three sets of training papers that trainees will use to practice applying the rubric for the prompt, and two to three sets of qualifying papers that will be used to certify that trainees are able to accurately apply the rubrics and proceed to score.

All papers, scores, and explanatory annotations will be forwarded to the CDE for approval prior to use. PEM scoring staff expect to work closely with the CDE in refining and revising training materials as needed to meet the needs of the CDE and the STAR Grades 4 and 7 Writing program.

Rangefinding Committee Meetings

PEM will design and facilitate rangefinding committee sessions to establish and document standards for scoring student responses from the field tests and the Operational administrations of the STAR Grades 4 and 7 Writing assessments. PEM's staff will work closely with the CDE to create an efficient and dynamic process that serves the needs of the STAR Program while providing a valuable professional development experience for California educators.

Rangefinding meetings will be held in California at a location to be determined in conjunction with the CDE and ETS. Committees will be composed of a representative group of California teachers and staff active in Grades 4 and 7 and in Writing. In general, there will be separate committees for the grade levels, with five to eight California teachers in each of these committees. Additional attendees will include CDE representatives, SBE test liaisons, ARP representatives, ETS test development representatives, PEM's Performance Scoring Project Manager and performance scoring directors assigned to each grade level to be reviewed.

Operational Rangefinding

For the operational prompts, student responses will be pulled from the original field tests and will be used for rangefinding meetings in advance of each operational administration. PEM assumes that ETS will make available the full set of student responses from the field-test administrations of any prompts included in the operational tests.

The operational committees will focus more intensively on the prompts to be included on the operational test forms for each grade level. The committees will set the standards for scoring these operational prompts by defining consensus scores on a large sampling of responses to each, which PEM scoring staff will use as the basis for developing scorer training.

Responses to "Crisis Papers"

Based upon CDE procedures, PEM scorers receive initial instruction regarding alerts during training. PEM's image-based scoring system provides a convenient method for alerting responses that may require local intervention, even if they are unsure whether intervention is required. PEM staff does not make that determination, but rather will forward any response in question to ETS for appropriate handling.

11. C6a. Scoring Grade 4 and Grade 7 CST Writing Tasks

The Image-Based Scoring System

Student responses are scanned from original test booklets, converted into an electronic format, and seamlessly distributed to the computer workstations of qualified, trained scorers.

The system automatically routes responses requiring second scores or resolution reads to qualified personnel.

All scores assigned to student responses are automatically captured and available for review.

PEM's image-based scoring system integrates multiple processes (routing work, scoring responses, monitoring quality, and tracking progress and workflow) into a single, efficient, user-friendly system.

11. C6a1. Criteria for Selection of Scorers, Materials Developed to Train Scorers, the Training Process, and Use of Electronic Technology for Scoring

Online Training

Prior to operational scoring, trainees will be required to complete a prompt-specific online training course containing multiple modules, including a qualifying test. Each module must be completed in sequential order, and trainees must pass the qualifying test to be certified for participation in scoring. There will be a unique course for each operational STAR Grade 4 and 7 Writing prompt. All of the student responses, corresponding scores, and explanatory annotations contained in the training modules will be approved by the CDE prior to use. The modules will include the following, and may be further customized to meet the needs of the STAR Grades 4 and 7 Writing program:

- Project overview
- Explanation of the Scoring Methodology.
- Scoring versus Grading
- Reader Bias
- Writing Prompt
- Rubric
- Anchor Set
- Practice Scoring
- Qualifying Sets
- Condition Codes and Alert Papers

11. C6a2. Scoring Student Responses that are Typed with a Computer or Communication Device

PEM will score computer or communication device-typed responses to the Grade 4 and 7 prompts. These documents will be flat-bed scanned for image capture, and then the images will be electronically routed for scoring, much like the responses written in the test booklets.

11. C6a3. Ensure validity and reliability of scores for the constructed-response items

Quality Assurance and Maintaining Reliability. In conjunction with PEM's ISO 9001 certification, PEM's Performance Scoring Center has developed and documented a standard system for addressing the complexities inherent in monitoring and maintaining quality throughout large-scale handscoring projects. PEM will maintain a quality assurance system that will consist of these components:

- Automated Backreading
- Blind Validity
- Targeted Calibration
- Messaging
- Reader Performance and Project Status Reports

11. C6a4. Methodology for Ensuring the Comparability of Scores for All Students Tested in Each Grade

PEM will utilize a blind validity system as well as calibration papers to provide comparability of scores for all students tested in each grade. The configurable blind validity system captures, calculates, and reports validity data. The validity feature is used to provide an objective and systematic check of scorer accuracy. In addition to the validity mechanism, calibration papers will be used to proactively promote accuracy. The calibration sets will provide supervision and guidance on how to deal with prompt-specific issues, score boundaries, or types of responses that are particularly challenging to score consistently.

11. C7. Optional Process for Requesting Verification of a Student's MC and/or Grade 4 or Grade 7 Writing Test Scores

Verification of Scores

As an additional measure for providing satisfaction with the scoring of constructed responses, PEM will conduct rescores for Grades 4 and 7 Writing prompts when parents/guardians, teachers, or administrators request it and this request meets the criteria put forth by the CDE. The original scores assigned to these student responses will be reviewed for accuracy by expert scorers.

PEM's Performance Scoring content experts will review each original student response in question, along with the original score assigned.

In each case, the original score assigned to the student response will be reviewed in close comparison to the original anchor papers used in training, and changed if warranted.

If PEM's expert reviewers determine that the original score assigned was incorrect, Performance Scoring will provide a new score. The new scores will subsequently be provided to the CDE.

11. C8. Analyses and Studies that will be Conducted to Ensure the Reliability of CAPA Scoring

CAPA tasks are scored using a 5-point rubric (Level I) or a 4-point (Levels II-V) holistic rubric approved by the CDE, which are designed to include specific behavioral descriptors for each score point to minimize subjectivity in the rating process and facilitate score comparability and reliability. Student performance on each task is scored by one primary examiner, usually the child's teacher, or by another licensed or certificated staff member who is familiar to the student and who has completed the CAPA training. To establish scoring reliability, approximately 10 percent of students receive a second independent rating by a trained observer who is also a licensed or certificated staff member.

To assure the reliability of CAPA scoring, student scores will be subjected to several types of reliability analyses:

- Internal Consistency Reliability
- Standard Error of Measurement
- Inter-Rater Reliability

The CAPA scores will also be analyzed exploring the methodology used for estimating the reliability of performance level classification decisions described in Livingston and Lewis (1995), and implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.12).

11. D. Analysis of Test Results for the CSTs, NRT, CAPA, CMA, and STS

11. D2. Analyses Necessary to Document the Reliability and Validity of Results for Individual Students

To verify the reliability and validity of student scores, the following analyses will be conducted:

- Verify that tests were built with item-total correlations that are sufficiently high to produce reliable scores at the student level, given the length of the test.
- Compute internal consistency reliability coefficients for major subpopulations to make sure that the tests are functioning similarly in each group.
- Compute classification consistency coefficients at all cut points.
- Conduct Differential Item Functioning (DIF) analyses for all field-test items to verify that the items do not favor one group over another. Groups who receive accommodations and/or modifications should be included in these analyses when sample sizes permit.
- Exclude C-DIF items from operational forms if they have not been subjected to review by panels with focal group representation.
- Longitudinal comparisons of data in the aggregate should be made at both early (just after equating) and late (at P1) points in the process to verify the reasonableness of the scores for comparable groups.
- Verify that the content of the constructed test forms meets blueprint specifications.
- Conduct a concurrent validity study using the 2004 CST and NRT data (This is the last year where NRT data existed for all grades).

11. D3. IRT Calibrating, Scaling, and Equating Procedures to Assure Comparability of Scores for the Duration of the Contract

To assure the comparability of the assessments from year to year, operational forms will be linked back to the base year through IRT scaling and true-score equating using a common-item equating design.

For all of the tests except CAPA and CST ELA Grades 4 and 7, all items are dichotomously scored and will be calibrated using the 1-parameter (1P) model.

The polytomously scored items in CST ELA Grades 4 and 7 and the CAPA test items (tasks) will be calibrated using the one-parameter partial credit (1PPC) model, a more restrictive model of the generalized partial-credit model (Muraki, 1992) where all items are assumed to be equally discriminating.

All IRT analyses will be conducted using the proprietary version of PARSCALE (Muraki, 1999) that is contained within the ETS GENASYS analysis package. In IRT-based equating, once the items from the two forms have been placed on the same IRT scale through their common items, raw scores on the new form are converted to raw scores on the reference form using the IRT parameter estimates for the respective tests. These converted raw scores are then transformed to scaled scores through a raw-to-scale table lookup and linear interpolation. For the CST and CAPA tests, "reference" scales already exist. For the new tests (CMA and STS), the first operational form, or set of forms, will establish the reference scale. It is important that the reference forms be built to the test blueprint to assure that the reference scale reflects the composition of the operational test. Otherwise, unintended biases may be established in the reference scale.

Although there are subtle differences in the analyses when tests are composed of dichotomously-scored items, polytomously-scored items, or a combination of both, the procedures to be used for equating the (CST, CMA, STS, or CAPA) tests involve the same three steps: 1) item calibration, 2) item parameter scaling, and 3) true score equating. These steps are described below.

11. D3a. Scaling the CSTs and CAPA

For the item calibrations, the PARSCALE program is constrained by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588). The resulting estimation is equivalent to the Rasch item parameter estimates.

11. D3b. Procedures for Developing and Verifying Accuracy of all Conversion Tables

Next, calibrations of the current administration data are scaled to the previously available reference scale estimates using the Stocking and Lord (1983) test characteristic curve (TCC) procedure. In the case of one-parameter model calibrations, this procedure is equivalent to setting the mean of the new form item parameter estimates for the common items equal to the mean of the previously scaled estimates.

Transformed new and old parameter estimates are evaluated using weighted (based on the reference form abilities) root mean square difference statistics that summarize differences in ICCs.

11. D3c. IRT Procedures for Developing an STS Scale and a CMA Scale for the First Administration

Once the new calibrations for each test are linked to the Rasch scale and defined by the reference calibrations, IRT true score equating procedures are utilized to transform the new form number-correct scores to their respective reference form scaled scores. The true score equating procedure is based on the relationship between raw scores and ability.

Equating samples will be selected from available student records obtained in the second cycle when the new CST form is administered. Only students with valid results on the respective tests will be included in the equating samples. The samples should be representative of the population of examinees with respect to key demographic and geographic factors, as well as disability classifications.

11. D4. Calibration, Scaling, and Equating of Braille Tests with the Print Versions

When all the items in braille tests can be brailled, the same scoring tables that are used for the regular tests are used. When one or more items in a test cannot be brailled, the test is shortened by this item and the shortened test is equated back to the reference form to produce a new scoring table.

11. D5. Steps, Procedures, and Software that the CDE can use to Replicate Calibration, Scaling, and Equating Procedures

ETS equating procedures can be replicated by an independent party or the CDE using either the commercial version of PARSCALE or WINSTEPS for item calibration. Other software may also be viable, although it has not been verified to date. The Stocking and Lord (1983) item-scaling procedure could be readily programmed or alternately, the average difference between the b-parameters for the linking items could be computed in a spreadsheet and applied to the raw calibrated item parameters. The procedure for finding the equated number correct score on the new form of the test could be readily programmed.

The student data used for the equating would be made available to the contractor conducting the replication as would the item parameter estimates (both scaled and unscaled), the scaling constant used, and the raw to-scale conversion tables.

12. Component Task 12: Reporting Test Results to LEAs (CST, CAPA, CMA, STS, NRT)

12. A. Expedited Return of Individual Student Results to School Districts

ETS will report student results (that is, scaled scores, performance levels, and cluster scores) via the Web before district and State aggregations are possible. In order for this expedited reporting of student scores to happen, ETS will introduce a Two-Form model of equating tests year-to-year.

The Two-Form Model

The two-form model avoids both the increased score variability inherent in pre-equating and the increased demand for early item development of the multiple operational forms model. ETS refers to this model as the "Two-Form" model.

The current method of STAR testing can be referred to as a One-Form Model since only one form of each CST is developed and used for each annual administration. In the Two-Form Model, the first part of the annual STAR administration window uses one test form while the second part uses a second test form. Additionally, the second test form for Year One (e.g., 2006) becomes the first test form for Year Two (e.g., 2007). As a result, the Two-Form Model does not require any more item development than the current One-Form model. Table 28, below, compares the timelines for the current One-Form Model with the recommended Two-Form Model:

Table 28. Timeline Comparisons of One-Form and Two-Form Models

Year and Month of Testing	Current STAR testing model (One-Form)	Recommended Two-Form model
Year 1, March-May	Form 1	Form 1
Year 1, June-July	Form 1	Form 2
Year 2, March-May	Form 2	Form 2
Year 2, June-July	Form 2	Form 3
Year 3, March-May	Form 3	Form 3
Year 3 June-July	Form 3	Form 4

Using this model:

- No additional item development effort is required in the Two-Form Model, compared to significant new development for the Multiple Operational Forms Option. The Pre-Equating Model would also require increased development for more field-test items. The Two-Form Model (post-equating design) is more efficient from a development perspective. Many fewer versions of the operational forms are required to field test new items, saving the State significant costs.
- The score variability inherent in the Pre-Equating Option is avoided. Because the equating is based on the psychometrically stronger procedure of post-equating, the State can place much greater confidence in the accuracy and stability of the equated scores of both the early and the later form.
- Expedited reporting of individual student scores is facilitated. Form 1 is already post-equated when it is administered to students in the first part of the annual STAR

administration cycle. Form 2 requires post-equating, but this equating procedure can proceed much faster because the second form is introduced at the beginning of a period when the largest number of students will be testing. More data coming in means that equating can proceed faster and thus equating tables and scoring can proceed faster. ETS estimates that the turnaround time for students testing in the May through August window will be only slightly longer for the majority of students than the turnaround time for students in the earlier part of the testing window (More information about turnaround time is presented later in this section).

- Score aggregations would be unaffected under the Two-Form Model since the two forms used in any one administration year are equated.
- Expedited reporting will begin in 2007.

Key Issues in Implementing the Two-Form Model

1. **When should Form 1 end and Form 2 begin to be administered?**
   The date of demarcation between the two forms is critical. If it is placed too early in the testing cycle, then there too few early testers who would benefit from the model. If it is placed too late, then there will be too few students and not enough data to conduct the equating on Form 2. ETS will work with the CDE to determine the optimum point for placing this date of demarcation. ETS recommends that the best way to determine this date is to work from the need to have sufficient data to conduct the equating of the second form each year. Typically, half of the student answer documents arrive for scanning by the second week of June. Half of the statewide data is sufficient to equate all but the smallest CST test volumes (for example, Integrated Math). This means that districts that typically test prior to mid-May would receive the first form for administration, while the remainder of districts would receive the second form for administration.
2. **How would districts with multiple administrations be handled?**
   Some school districts administer the STAR Program at more than one time during the calendar year depending on the calendars of specific schools. In some cases, one administration might occur prior to the date of demarcation while the others might occur following that date. ETS recommends that such districts all be administered the second form, even though some students within those districts would take Form 2 prior to the date of demarcation for the State.
3. **Which CSTs are amenable to expedited reporting?**
   ETS will expedite reporting of CST scaled scores and cluster scaled scores. However, not all CSTs are subject to expedited reporting. For example, Grades 4 and 7 English-Language Arts (ELA) CSTs require an essay. The essay is given in one of two fixed-date windows during the year. While it is possible to report ELA scaled scores without the essay component, this would result in possible inconsistencies in reporting. Additionally, some CSTs are administered to fewer than 5,000 students statewide each year. Tests such as the Integrated Math and Science tests might not present sufficient student scores by the date of demarcation to allow for equating of the second form in the same time frame as the more voluminous tests. Working with the CDE, ETS will determine the most feasible combination of individual student scores for expedited reporting.
4. **Will all student test scores be part of expedited reporting?**
   Because of the use of the Two-Form model, all multiple-choice test scores including low-volume end-of-course tests will be expedited. ELA scores for Grades 4 and 7 include essay scores. Essay scores are not available at the same time that multiple choice ELA scores are. There are three ways to address Grade 4 and 7 ELA scores in expedited reporting.

a. Do not report them. Districts would still get these scores on paper reports delivered each summer.

b. Report ELA scores at Grades 4 and 7 without the essay scores factored in. The STAR scoring system has already developed a method for deriving scaled scores without the essay to handle cases of missing or mismatched essays

c. Add Grade 4 and 7 ELA scores to the expedited score reporting record after student essay scores are available, perhaps several weeks after the rest of students' scores are posted but before annual equating is completed.

ETS will work with the CDE to determine which of these options should be implemented.

5. How and how quickly would expedited student scores be reported?
ETS will expedite the reporting of individual student scores by means of a secure Web site that is accessible to district STAR coordinators. This process supplements the reporting of student and aggregated score reports on paper. The paper reports for student, school, district, county and State reports will also be accelerated as described below in this section.

Following is a brief timeline describing how expedited student scores would be reported for students whose districts used the first test form in an administration year:

Table 29. Brief Timeline of Reporting Expedited Student Scores using Form One

Days After Receipt of Answer Documents	Process
0	Receipt of student answer document
1–10	Quality control and scanning
11	Student Data scored
12–14	Student data posted to secure Web site

The score data provided for each student would consist of scaled scores for all CSTs with the exception of Grade 4 or 7 ELA scores (due to lack of essay) and, possibly, low-volume CSTs such as Integrated Math and Science. Depending on analysis of volumes of answer documents, expedited student reporting for STS, CAPA, or the CMA could also be provided.

For districts using the second test form in a given year, there is an additional consideration required: the actual equating process is occurring in mid-June. This adds approximately 10 days to the process of expedited reporting for those districts at the very front of the Form 2 administration window. Taking this into consideration, the following timeline is likely for districts testing during the first two weeks of the Form 2 administration window:

Table 30. Brief Timeline of Reporting Expedited Students Scores using Form Two

Days after Receipt of Answer Documents	Process
0	Receipt of student answer document
1–10	Quality control and scanning
10–20	Equating of Form 1 and Form 2
21–23	Student Data scored
24	Student data posted to secure Web site

For the majority of students tested with Form 2, the 14 calendar-day cycle for Form 1 students will still apply.

It is important to note that the reporting process can only be expedited once answer documents are received. ETS will make every effort to remind and encourage districts to return their answer documents within the required five-day window following the last day of their testing window. ETS will also reach out to districts to assist them in returning materials as required. However, ultimately it is the responsibility of the districts to return answer documents in a timely manner if expedited reporting is to have its greatest effect.

To support the expedited posting of student results online, PEM will deliver student scores to ETS within 14 calendar days of receipt of student answer documents. ETS will be able to post student results for first-form testers online almost instantaneously. Student results for second-form testers will need to be equated by ETS first, but results will be available online much earlier than when paper reports are currently received.

In order for this Web-based expediting of student scores to occur, additional assumptions about expedited reporting of student results are:

- ETS, PEM, and the CDE will work together to identify a cut-off date for first-form testers, whereby districts with test windows that start after that cut-off date will take the second form. In order to minimize the time between receipt of scorable materials at PEM and the beginning of equating for the second form, ETS assumes a cut-off date of May 15.
- ETS will provide PEM with two sets of conversion tables for each test when the Two-Form model applies. The Two-Form Model will need to be phased in for STS and for the CMA as it becomes operational.
- At Grades 4 and 7 the student files will contain multiple-choice scores only. Grade 4 and 7 Writing will not be included in the expedited posting of student results.
- For those tests where students are assessed on more than one answer document (for example, grade level and end-of-course), PEM will match the scores using the criteria identified in Section 7.
- Changes made by districts during the demographic edit process will not be included in the expedited posting of student results. PEM will resend student files to ETS with the writing results and demographic changes for refreshing online scores once that information is available.

Improving Reporting through Pilot Projects

Expediting Paper Score Reports

Although not amenable to the same quick turn-around as the individual student scores via the Web, ETS and PEM are also introducing procedures that will speed up the return of paper reports, both individual student reports as well as aggregate reports to districts.

ETS and PEM will streamline and consolidate the production and generation of all paper-based reports into a single process. Currently, paper-based reports are generated via four separate processes. Student record labels and all summary-level reports are generated, printed, packaged, and distributed as a group. Student reports are then generated, printed, packaged, and distributed in two different groups; one for color reports and one for black-and-white reports. Teacher reports are generated and distributed on a fourth later schedule from all other paper reports.

ETS and PEM will combine all paper-based reporting into a single process to deliver all reports in a single consolidated process that will align with the schedule of current STAR Program for

those reports that are delivered first. Of most importance and value to the end users of these reports, the districts, schools, and teachers, will be the ease of use and ability to compare the entire suite of reports produced at a single time. The logistical requirements of having to sort and distribute the STAR reports within the districts will be reduced to a single event thereby significantly reducing burden on district and school staff and increasing the immediate usability of all reports. Questions that may arise regarding the reports will be more readily resolved within the district and those that cannot be resolved will be able to be addressed with all reporting components available for review together rather than having to wait for disparate reporting processes to be coordinated.

The ability to speed reporting into a single administration will be accomplished due to a number of factors including:

- Reduction of data handoffs between systems and organizations through the entire reporting process;
- Improved quality control due to a reduction of versions or instances of the same sets of data in the process, leading to less opportunity for the introduction of discrepancies in the various components of the reporting processes;
- Reduction in duplication of effort and cost in generating datasets from which reports are generated, printed, and distributed;
- Teacher reports will be delivered with the student and aggregate reports. ETS and PEM propose removing State-level data from the Teacher Report, thus removing the current impediment to prompt delivery of those reports.
- Delivery of student reports fully two weeks earlier than is currently achieved as a result of these process improvements.

The single greatest benefit to the districts and schools may well be the delivery of a single integrated, well-designed package of reports which will greatly increase their immediate usability and streamline the distribution efforts within the districts and schools and to the parents of each student. This coordinated effort should also allow the CDE, districts, schools, and teachers to more effectively and efficiently respond to questions since they will have the full set of reports available to them at once.

Reducing Turnaround Time of Scores via Computer-based Testing Pilots for CST

(Currently in Section 12.A, ETS recommends moving to Section 10 "Delivery and Collection of Test Material")

In order to expedite return of student test data and scores, ETS will plan for implementation of Computer-based Testing (CBT) of CSTs starting at the high school level. Although a full implementation of CBT for all CSTs may not occur until after the current contract, ETS will provide the test development, delivery, and scoring infrastructure scalable to support a full implementation of CBT for CSTs.

To prepare for CBT implementation, ETS will conduct an extensive survey of district and school readiness to implement CBT, and will provide an analysis of what other states are doing with computer-based testing. This survey will be designed, approved, conducted and analyzed between March and July 2006. There will be both a survey instrument and visits to a number of districts to be specified by SBE staff and liaisons and the CDE to assess readiness for CBT implementation. Issues to be addressed in the survey will be determined by CDE and SBE staff and liaisons and may include:

- Existing firewalls, computer hardware, including network, number of computers, accessibility for testing

- Telecommunications infrastructure
- Onsite technical assistance to handle technical problems
- School and district personnel willingness and expertise to implement CBT

Based on the results of this survey and related onsite visits, ETS will prepare a report for the SBE and CDE with recommendations on the rate of rollout of CBT.

The sequence and extent of rolling out CBT for each CST will be determined by SBE staff and liaisons and the CDE. The project will begin in the summer of 2006 with the completion of a School District Infrastructure Survey. ETS will also present the CDE and SBE staff with an analysis of what is being done in other states in terms of CBT. Both reports will be presented to CDE and SBE staff for review by December of 2006. ETS will provide a "minimum requirements" document for CBT equipment and district readiness by December of 2006 that would be needed to assist districts in implementing CBT.

To accomplish a statewide implementation of a low-volume CBT pilot for the 2008 test administration, ETS would need approval of a pilot test from SBE staff and testing liaisons and the CDE no later than January 1 of 2007. Additional costs to complete CBT pilot activities are contingent upon approval by CDE and SBE of a contract amendment and upon sufficient funds being made available by the Legislature in future fiscal years.

Expediting the return of results from low-volume tests will speed up the entire paper score reporting process since they take the most time before enough answer documents are returned to allow for annual equating. For planning and budgeting, ETS has assumed Integrated Math 1, 2, and 3, and Physics CSTs would be available via computer in 2008. Each of the Math tests had fewer than 10,000 test-takers statewide in 2005. Physics had approximately 60,000 test-takers in 2005. The final selection of CSTs for 2008 implementation will be made by SBE staff and liaisons and the CDE by January 1, 2007.

In 2008, students taking these CSTs will either be administered a CBT version of the test or the standard paper-based version. This will allow ETS to address comparability between the two modes of administration. In subsequent years, a paper-based version will be available, but the majority of students taking these CSTs will do so via CBT.

ETS will work with the CDE to establish operational requirements for the CBT infrastructure including form creation, item display, security requirements, data capture and scoring.

For each CST delivered via CBT, test items will be displayed in a way similar to how they appear in the paper test booklets. All CBT versions will be reviewed by the appropriate ARP prior to implementation, with ARP recommendations informing the decision to implement CBT for a given test.

The method of delivering CBT to schools will be "Internet-based Testing" (iBT). This method will minimize the need for schools to have specialized computer laboratories to provide CBT. Starting in July 2007, the STAR order management system will treat CBT for a given CST as a separate testing order. To order CBT tests, a district must first be approved by CDE for CBT. The criteria for such approval will be determined by SBE staff and liaisons and the CDE.

ETS will only begin work on the 2008 CBT pilot after SBE staff and liaisons and the CDE have reviewed and approved the pilot test. By September 2008, ETS will provide SBE staff and liaisons and the CDE with a detailed scope of work, timeline, and budget for the 2009 CBT administration. Additional costs to complete the activities are contingent upon approval by CDE and SBE of a contract amendment and upon sufficient funds being made available by the Legislature in future fiscal years.

Reducing Turnaround Time for CAPA via Web-based Entry of Examiner Ratings

CAPA is low-volume test. District test results cannot be reported until sufficient CAPA answer documents (rating sheets) have been returned to allow for equating with the previous-year's test. To expedite the return of CAPA Examiner ratings and the scoring of the CAPA test, ETS will make available a secure Web site for CAPA Examiners to enter their ratings online. This Web site will allow an examiner to identify the student being administered the CAPA test, enter their ratings online and submit them directly to ETS. ETS will then process those ratings as they would ratings submitted on CAPA answer documents. CAPA examiners will be trained to use this Web site during the 2007 Pre-Test Workshops. Additional training will be provided through the COE STAR liaisons. All CAPA answer documents will still be returned to ETS. ETS will review a sample of such answer documents to check the proper receipt of Web-entered examiner ratings.

12. B. Overall Reports

12. B1 Preparation of Reporting System

PEM will create detailed report test plans based on the agreed-upon customer requirements. The report test plan consists of test cases representing varying combinations of districts, schools and grades. Each of the test cases is structured to produce a specific circumstance that will be validated on the reports. Test cases are reviewed and augmented for each test administration to verify that software system conditions are accurate.

For report quality control, four general areas are evaluated including:

- Comparing report formats to input sources from the CDE-approved samples
- Validating and verifying the report data by querying the appropriate student data
- Evaluating the production print execution performance by comparing the number of report copies, sequence of report order and offset characteristics to the CDE's requirements
- Proofreading of reports at PEM prior to any district mailings and then sending of the reports to the ETS and the CDE for review

All reports will be required to include a single, accurate CDS code, a charter school number if applicable, a district name, and a school name. All elements will conform to the CDE's official CDS code and name records. From the start of processing through scoring and reporting, the CDS Master File is used to verify and confirm accurate codes and names. PEM receives updated CDS Master Files from the CDE throughout the year.

For students assessed on more than one answer document, the matching process, as described in Sections 7 and 10 will provide for the creation of individual student records from which reports will be created.

After the reports are validated against the CDE's requirements, a set of reports for pilot districts will be provided to the CDE and ETS for review and approval. PEM will send paper reports on report form, foldered as they will look in production. The CDE and ETS will review and sign-off on the report package within five to seven business days.

Upon the CDE's approval of the reports generated from the system test, PEM will proceed with the first production batch test. The first production batch is selected to validate a subset of districts that contain examples of key reporting characteristics representative of the State as a whole. The first production batch test incorporates customer-selected school districts and provides the last check prior to generating all reports and mailing them to the districts.

12. B2 Report Production/Printing

PEM will:

- Print each page as original, thus producing easy-to-read reports that do not smudge
- Utilize a sophisticated report collation process combined with high-speed laser printing technologies to print all report types in continuous print streams

PEM will print, package, and distribute the following reports for the STAR Program.

- Teacher Report
- Student Report
- Student Record Label
- Student Roster with Results, by Grade (Master List)
- Summary, by Grade (Master List Summary)
- Subgroup Summary, by Grade
- Unmatched Reports
- Electronic Student Data File

The Materials List described in Section 9. D2 breaks out the reports by test type, and provides assumptions regarding the number of copies per report, the level of aggregation and the report recipient. In addition to the number of copies per report cited in the Materials List, districts can order additional copies of reports on the Processing Form when documents are returned for scoring. PEM will provide a reprint capability for all reports to support such requests.

12. B3 Report Packaging

Each shipment of reports for schools, districts, and counties will include a specific letter enclosed with the package describing what they are receiving in their shipment. All reports will be assembled by grade, school, and district. School sets of reports will be assembled and shipped to the district for distribution to schools.

For those districts that test in multiple test administration windows, report packages for the early administrations will include:

- Student Data File
- Student Reports
- Student Record Labels
- Student Roster with Results (Master List)

For those districts that test in a single test administration window, and for the last administration of a multiple test administration window district, report packages will include:

- Student Data File
- Student Reports
- Teacher Reports
- Student Record Labels
- Student Roster with Results (Master List)
- Unmatched Reports

- Summary (Master List Summary)
- Subgroup Summary

CST, NRT, CAPA, STS, and CMA results will be incorporated into the Student Data File, Student Roster, and Subgroup Summary reports. Separate Summary reports will be provided for end-of-course tests and the NRT.

The following packaging processes will be followed to provide districts with clearly organized shipments:

1. All reports will be assembled and placed in report folders. One color of folder will be used for district reports; another color will be used for building reports. Assembly clerks will check a district's printed reports against the internal reports packing list, before placing the reports in folders. Confidential student reports will be grouped in accordance with the information provided on the packing list, and placed in folders.
2. Reports will be boxed and labeled by school with the boxes for all schools within each district shipped to the district STAR coordinator for distribution within the district. The district reports will be boxed separately.
3. Enclosed in each shipment of reports will be a specific letter describing what the district is receiving in the shipment.
4. A pallet map will be included with each report shipment for districts that receive more than one pallet of reports.
5. Prior to shipment, quality control specialists will perform a final quality check of reports and check for complete units of work, correct assembly, and the correct use of mailing labels.

12. B4 Report Delivery

PEM will work with ETS and the CDE to design reports and reporting systems that provide accurate results to all stakeholders in a timely manner.

PEM will distribute all paper reports so that they are received by districts no later than August 8 for district receipts received by July 1. All reports for districts authorized to test during July and August will be received by the districts no more than five weeks after receipt of answer documents at PEM.

For those districts that test grade levels in multiple testing windows/administrations, they will receive individual student reports as processing and scoring is completed for each administration. Summary reports for the multi-administration districts will be received within approximately eight weeks of the delivery of the last administration to PEM or by August 8; whichever is later.

PEM will implement the following improvements to the current reporting system:

- Student and Teacher reports will be shipped with the rest of the districts' report shipments, and not as separate shipments, which is currently the case.
- Both copies of the student report will be originals. Currently one copy is in color and the other copy is black and white.
- All report designs will be improved to include features that have been communicated to be important to the end users. These enhancements are further described in Section 12. C.

A single entity (PEM) will be responsible for all report production and distribution, eliminating a series of hand-offs that currently exist with the reporting vendor for Student and Teacher reports.

Box 1 of each district shipment, which contains the letter explaining what is included in the shipment and the summary reports for the district, will be white so that it will be easy for the district STAR coordinator to easily distinguish this box from other boxes in the shipment.

Trained shipping personnel will determine the most reliable and rapid means of delivering each shipment of reports. Each district's reports will be entered in the shipping manifest system as they are shipped. PEM's barcode technology combined with distribution partners' (UPS, for example) tracking systems will allow PEM to provide instant updates about the location and status of report packages should any problems arise. Upon receipt of reports at the district, district personnel signatures will be required to provide for secure delivery.

PEM will transmit status records to ETS at the following points, so that they are also able to track where a district is at a given point in the reporting process:

- Reports printed
- Reports shipped
- District complete

12. B5 Synergy with the CSU EAP Program

ETS proposes to coordinate between the STAR Program and the CSU Early Assessment Program (CSU-EAP) to provide timely reporting of EAP results to districts and students to the extent that such coordination does not negatively affect the execution of the STAR contract. Proposed activities will begin in the 2007–08 school year. ETS will provide proposals on the following activities to CDE and SBE staff for review and approval:

- Use of STAR Pre-ID data to provide parent address information for CSU EAP student reports;
- Posting of CSU-EAP student status data in aggregated form on the STAR results Web site;
- Investigation of incorporating CSU-EAP results on STAR parent reports;
- Coordination of CSU-EAP test development activities with STAR test development activities to facilitate the printing and shipping of STAR testing materials;
- Coordination of the Grade 11 CSU-EAP essay testing window with STAR Grades 4 and 7 testing windows to facilitate return of CSU-EAP essays for quicker scoring and reporting.

The above activities may necessitate an inter-agency agreement between CDE and CSU. ETS will not proceed with any activities requiring an inter-agency agreement until the written agreement is secured.

All such activities must first be reviewed and, where necessary, approved by SBE liaisons and staff and the CDE before being implemented. In no case will the STAR Program bear the contractor costs for such activities. In addition, in no case will such activities impede the administration, scoring, and/or timely reporting of STAR results or the fulfillment of any obligations under the STAR contract. If such activities jeopardize the timeline of the STAR contract, STAR-funded activities will be discontinued.

12. C. Production and Distribution of Paper Score Reports

Table 31 provides a matrix of the reports by test type, and provides assumptions regarding the number of copies per report, the level of aggregation, and the report recipient.

Table 31. Report Matrix

Report	No. Copies	School	District	County	State	Aggregation Level
Student Report						
Traditional Schools	2*	1	1			
Independent Charter	2*	2				
California Report for Teachers	1	1				
Student Record Label	1	1				
Student Master List	1	1				
Group Summary CAT/6						
School	2	1	1			School
Independent Charter	2	1		1		School
District	2		1	1		District
County	1			1	1***	County
State					1	State
Student Master List Summary — CST for each grade level 2–11						
School	2	1	1			School
Independent Charter	2			1		School
District	2	1	1	1		District
County	2			1	1***	County
State	1				1	State
Student Master List Summary — End-of-Course CSTs for each math test grades 8–11 and science test for grades 9–11						
School	2	1	1			School
Independent Charter	2	1		1		School
District	2		1	1		District
County	2		1	1	1***	County
State	1			1		State
Subgroup Summary						
Reports for the following Subgroups: • Male/Female • Disability/Not Disabled • Econ. Disadvantaged/Not Econ. Disadvantaged • English Learners/EO, I-FEP, and R-FEP • English learners in CA Public School less than 12 months • English learners in CA Public Schools 12 Months or More						
School	2	1	1			School
Independent Charter	2	1		1		School
District	2		1	1		District
County	2			1	1***	County
State	1				1	State
Student Data File						
School Electronic Student Data File	1	1				
Independent Charter						
District Electronic Student Data File	1		1	1**		
State Electronic Student Data File	1				1	

* 1 copy for parent/guardian and 1 copy for current teacher

** County Offices of Education only for schools operated by the county

NOTE: Independent Charters are treated as districts for aggregation purposes.

*** State copies of county reports are to be electronic copies on CD-ROM

In addition to providing student and summary performance data, ETS and PEM will provide unmatched reports to districts for resolution.

These reports will be sent with the district's other reports, and will include:

- An alphabetical listing of students within grade, within school, within district for whom there are multiple-choice scores but no writing test scores, and vice versa (Grades 4 and 7)
- An alphabetical listing of students within Grade 3, within school, within district for whom there are CST scores but no NRT scores, and vice versa
- An alphabetical listing of students within grade, within school, within district for whom there are CST end-of-course scores but no grade level scores
- An alphabetical listing of students within grade, within school, within district for whom there are STS end-of-course Math scores but no grade level scores

Sections 12. C1 and C2 below describe the plans for producing SBE-approved reports, including proposed design enhancements resulting from PEM's research study and ETS's prior experience with design of the STAR student and teacher reports.

12. C1. Student Reports

PEM will design and produce the following types of individual student reports for the STAR Program:

- CST/NRT
- CAPA
- STS
- CMA

All student reports, except for the Grade 8 through 11 CST reports will be designed to print on both sides of a single sheet of 8.5 inch x 11 inch paper. Due to the number of subject areas assessed at the higher grade levels, including end-of-course assessments, the Grade 8 through 11 CST reports will be designed to be printed on both sides of a single sheet of 11 inch x 17 inch paper, which when folded becomes four 8.5 inch x 11 inch pages, or a report folder.

The proposed student reports include all of the same data points that are in the current student reports, including: scaled score for ELA, scaled score for Math, percent correct by content area, percent correct by content area compared to State Proficient range, and national percentile rank for NRT (Grades 3 and 7 only). Information on additional resources will also be provided.

Enhancements to the student report include the following:

- Better use of space improves overall viewing and ability to understand report.
- Identifying information such as student name is more predominant.
- Performance by content area is presented in bar chart format, which is more easily understood by parents.

In addition to percent correct, the number correct for each content area is provided.

Instead of describing all ELA and Math content standards, only those content areas for which the student needs to focus on are included to the right of the bar chart.

The student address was moved to the left side of the report to accommodate the use of left-windowed envelopes by districts.

A POSTNET barcode was added above the student address. The barcodes will be printed with the addresses by PEM, and will allow districts to qualify for postage lower rates and take advantage of faster, more efficient mail processing.

The STS Student Report will follow the same content and design rules used to create the CST Student Report, but will be in Spanish except for the California content standards in ELA and Mathematics, which will not be translated.

PEM will provide two copies of student reports to each district. One copy will be packaged for the district and the second copy for the school at which each student was tested. Both copies of the report will be provided in color. Color schemes will be chosen so that the district report can be easily photocopied as a black and white copy.

An adhesive Student Record Label for each student will also be provided to each district, packaged by school with the reports. The student record labels are printed five per sheet, one label per student, and include the student's overall test results. To comply with Section 60607(s) of the California Education Code, schools are required to affix this label to the individual student's permanent school records.

Student Record Labels will include:

- Student identifying information, such as student name, grade enrolled, grade tested, test date, date of birth, school and district where the test was taken
- California Reading List Number
- Accommodations and/or Modifications, if applicable
- Tests and subjects taken
- Scaled score
- Performance level
- National Percentile (NRT only)
- National NCE (NRT only)

The design of the student report will retain the look and feel of the current STAR student report. It will incorporate feedback from districts, SBE, its staff and liaisons, and CDE as part of the annual cycle of report specifications. Any changes to this report agreed to by December 31 of a given year will be incorporated into the following year's student report.

12. C2. Summary Reports

Under ETS's direction, PEM will produce and disseminate the summary reports listed below.

Below is a description of each report.

Student Roster with Results, by Grade (Master List). One copy of this report will be produced for each school, to provide school administrators with a single list of all students and their scores for a grade, or year-round schedule within a grade at a school.

Student names will be printed in alphabetical order within each grade, by last name, first name, and middle initial. Each student's CST and NRT scores will be printed. If scores are not available, a reason code will be printed. Subsequent lists will be printed that include the names of all students assessed with CAPA, CMA, and/or STS, as the programs are

phased in and if there were any students assessed with these tests at the school. These lists will also be in alphabetical order by student last name.

The report will print in landscape versus portrait format, and included "star icons" to reflect student performance proficiency levels visually, so school administrators can easily scan the report and get a picture of individual student performance. Key data such as student information, use of accommodations or modifications, scaled score, performance level, reporting cluster percent correct, writing applications standards score for grades 4 and 7, and national percentile and national normal curve equivalent for the NRT grades 3 and 7 will be provided.

Summary by grade (Master List Summary). The Summary by grade report will be provided at the following levels of aggregation: school, independent charter, district, county, and State. It will summarize performance of a group of students within a grade for the aggregate level, on the CSTs, CAPA, CMA, and STS. Separate summary reports will be provided for the same aggregate levels for the end-of-course tests and for the NRT.

Like the Student Roster with Results report, the Summary by grade report has been redesigned to be printed in portrait versus landscape orientation.

In addition, the following enhancements were made to the report design:

- A pie chart has been added to illustrate percent of students Proficient at each performance level for the level of aggregation.
- Bar charts have been added to visually depict student performance by reporting cluster.
- Writing assessment information has been presented in an easy-to-read format.

The end-of-course and NRT Summary reports will follow the same design contents, where appropriate. For all Summary reports, the number of copies required in the STAR RFS will be provided.

Subgroup Summary, by grade. A subgroup summary by grade at the following levels of aggregation will be provided:

- School
- Independent charter
- District
- County
- State

The report will allow schools and districts to look at results based on the following demographics:

- Disability status
- Economic status
- Gender
- English-language fluency
- Primary ethnicity

The report will be sorted by subgroup in the order specified by the CDE.

Overall performance levels will be broken down by specific demographics for the CSTs, the NRT, CAPA, STS, and CMA tests. PEM will work with ETS and the CDE to provide a Subgroup Summary report design that has the same look and feel as the other revised reports, while maintaining the content of the report.

In addition to providing hard copies of all reports, PEM will provide to CDE an electronic file via a CD-ROM for the following reports:

- Summary, by grade reports (County and State)
- Subgroup Summary, by grade reports (County and State)
- Electronic student data files are described in Section 12. E.

All STAR summary reports will be produced and delivered according to the following requirements:

- School-level reports will be produced by grade and will include the number of students tested, the number of valid scores, the average scaled score, the standard deviation of the scaled score, and normative data for the NRT.
- District summary reports will be comparable to the school-level reports. Districts will receive summary reports no later than August 8 of each year.
- County-level reports comparable to the school and district reports will be produced. County offices of education will receive county-level reports no later than August 8 of each year.
- A State-level report comparable to the school, district, and county reports will be produced. The CDE will receive State-level reports no later than August 8 of each year.

12. C3. Plan for Distributing Reports to LEAs

Reports will be distributed to districts, independent charters, counties, and the State. District shipments will include school reports, boxed separately by school. Each report shipment will be entered in the shipping manifest system as it is shipped. PEM will use barcode technology to provide instant updates about the location and status of report packages should any problems arise. Upon receipt of reports at the district, a district signature will be required to provide for secure delivery.

12. C4. Printing, Packaging, and Distributing Reports

PEM will print, package, and distribute the reports listed in this section.

12. C5. Identifying Charter Schools

As changes to school and district information arise, the CDE will update the CDS Master File and provide an updated file to ETS and PEM. PEM will use the CDS Master File for the start of processing through scoring and reporting to verify and confirm that accurate codes, names, and addresses are being used.

12. C6. Procedure for LEAs to Notify ETS that Complete and Accurate Reports/Files Were or Were Not Received

Each shipment of reports for schools, districts, and counties will include a letter describing how the reports are packaged to assist the district with report distribution. All reports will be assembled according to grade, school, and district. School sets of reports will be assembled and shipped to the district for distribution to schools.

If for any reason, a district receives what they believe to be incomplete or inaccurate reports or files, these procedures will be followed:

1. The district contacts STAR TAC to report the potential error.

2. STAR TAC requests that PEM investigate the potential error.
3. If it is found that ETS or PEM have caused an error, PEM produces revised reports and/or files, clearly indicating "revised" with the revision date on the reports, and ships the revised materials to the district.
4. ETS submits a report to the CDE and the district detailing the resolution of each inaccurate or incomplete report.
5. If the error is caused by ETS or PEM and district notification is received no later than July 26, it will be corrected for the August student data file and Internet Report posting.
6. Aggregate or student data files found to be incorrect due to ETS or PEM error will be corrected and reissued to the CDE at no additional cost.

12. D. The California Report for Teachers

The California Report for Teachers will be provided for STS and CMA components of STAR as well as for CAPA and CST components. For each type of test, separate reports will be provided for ELA and Math. Reports will be printed for individual teachers if:

- Teacher names and course periods were included for students in the district Pre-ID file; or
- If student answer documents were packaged for scoring by teacher and the teacher's name was gridded on the School and Grade ID sheet.

The current Teacher Report format will be reviewed annually and presented via focus groups to educators to inform the report specifications for the next year's report.

Reports will be printed by grade level or math course if student answer documents were submitted for scoring by grade.

To expedite the return of Teacher Reports, PEM will produce the Teacher Reports, package them by school, and ship them with the other student and summary reports to the district. Under the contract, PEM, as subcontractor to ETS, will produce all reports, including the Student Report and Teacher Report, and will ship all reports as complete shipments, reducing the turnaround time required for providing the reports to the districts.

Reports for individual teachers will be in sealed envelopes addressed to the teachers. Other Teacher Reports (for a grade level or a math course) will not be in envelopes, but will be part of the school's reporting package. It will be up to school administrators to determine the appropriate distribution of those Teacher Reports within their school.

The design of the Teacher Report will retain the look and feel of the current STAR Teacher Report. It will incorporate feedback from districts, SBE, its staff and liaisons, and CDE as part of the annual cycle of report specifications. Any changes to this report agreed to by December 31 of a given year will be incorporated into the following year's teacher report.

12. E. Electronic Student Data Files

In addition to the paper reports, electronic student data files on CD-ROM will also be provided for:

- Districts
- County Office of Educations, for those schools operated by the county
- Independent charter schools

Student data for all STAR operational tests (CSTs, NRT, CAPA, CMA, and STS) will be provided on a single data file versus separate data files, as space permits. The sort order will

be by CDS and by test. Electronic student data files will be sent with other report packages, such as student and summary reports.

The electronic report package for the State will include:

- Summary, by grade (also provided for end-of-course and NRT)
- Subgroup Summary, by grade
- PDFs of all county Summary reports

12. F. Interpretation Guidelines

ETS will produce interpretation guides for both CST and CAPA Student reports. Additional reports and interpretative guidelines will be added for the STS in 2007 and for the CMA when it becomes operational in 2008.

Although the reports themselves are self-explanatory, ETS will have most parts of the reports translated and will offer further definitions. The language will be as simple and straightforward as possible. The report will be one page, two-sided. It will begin with a statement of purpose with caveats and end with a short glossary.

These parts of the report will be directly translated:

- Superintendent's letter
- How to use
- California Reading List
- NRT
- Lists of resources

In addition, these parts of the report will be explicated:

- Student identification
- Scores and performance levels
- Content areas
- Information on standards
- NRT scores
- Resources

After the CDE has approved the wording of the interpretative guidelines, they will be formatted in print-friendly MS-Word, PDF, and HTML and posted on startest.org.

12. F2. Provide Guidelines in Multiple Formats in the State's Six Major Languages in Addition to English

Because the text of the reports varies so, not only from student to student, but also by grade and number of tests, it is impossible to translate every part of the guides.

Instead, three guidelines will be translated for CSTs, CAPA, and CMA into the following languages:

- Spanish
- Vietnamese
- Hmong

- Cantonese
- Tagalog/Pilipino
- Korean

For STS, two interpretative guidelines will be available in English and in Spanish.

San Joaquin County Office of Education (SJCOE) will handle the translations as they have for other districts, contractors, and for the CDE.

Each translated guideline will display the language of the translation in English prominently on the front. Each translation will also be formatted into MS-Word, PDF, and HTML and posted on startest.org. Both the English version and each translation will be printed so that ETS can send a paper copy to each district and site with student reports. This way, the district or site may copy the guides and translations in the appropriate amount and distribute them to parents as they choose.

12. G. Reporting and Correcting Errors

Any time there are changes to the data that require reports to be reprinted, STAR Technical Assistance Center staff will call PEM to print and ship revised reports. All of the reports will be clearly identified as “revised” with the appropriate revision date.

Reporting Errors

In any such event, ETS staff will take the following steps:

- Initially analyze the situation
- Inform the CDE immediately
- Further analyze the impact of the error
- Discuss solution options with the CDE
- Deliver an expedient resolution that best mitigates program risk.

Demographic Data Correction Opportunities for Districts

Districts have a set of opportunities to submit and correct accurate demographic data throughout the test administration life cycle.

Months before the administration, districts have the opportunity to submit Pre-ID files to CSIS and then to ETS for STAR Pre-ID processing. CSIS and ETS error summary reports will inform districts of omissions, errors and deficiencies and allow opportunity for file resubmission or online corrections.

Critical Pre-ID and gridding deficiencies are corrected during the Scan Edit resolution process (see Section 11. B). These deficiencies are related to CDE regulations about missing data for certain key fields exceeding tolerance thresholds.

ETS will offer a new data correction window immediately after online test results are posted for a district’s administration (see Section 12. A) and prior to paper report production and delivery for the district. This allows the district to correct errors prior to paper and Internet reporting. This error correction will improve the accuracy of the paper and initial Internet reporting (P1) and assure that districts receive the appropriate AYP apportionments. It also allows the opportunity to update student addresses in order to improve the accuracy of report mailings.

Another data correction window will be offered after all districts’ results have both been posted to the Web and reported to the districts on paper in mid-August. This allows the correction of additional demographic data omissions or errors that get identified from the Web or paper

reporting. This correction window takes place from late September through early November following each year's administration.

A final data corrections capability for districts will be offered after the final Web posting of summary results. This Post-P3 data correction window allows districts to correct data that was not previously identified as being in error. ETS sends the corrected data to the CDE for AYP/API consideration. Error Correction Opportunities summarized error correction opportunities during the STAR testing cycle.

Table 32. Error Correction Opportunities

Reporting of Demographic Data	Correction Date	Subsequent Demographic Data Correction Capability
Pre-ID submission error summary report	November through July, prior to administration	District STAR coordinators can resubmit Pre-ID file or add/correct data online; districts can also request expedited labels just prior to test administration for new students or to correct previously submitted data.
Scan Edit Resolution Report, detailing missing demographic data in excess required thresholds (see Section 11..B)	March through August, after administration	Districts can submit updates within a 48-hour window, prior to scoring.
Expedited online student-level results posting to districts	March through August, after the administration	New expedited data correction window is applied to paper reports and P1/P2 Internet Summary Results Posting.
Paper Reports, P1/P2 Web Summary Results Posting	September through November, after the administration	Currently-offered data corrections window to be applied to P3 Web summary results posting and delivery of student level data to the CDE
P3 Web Summary Results Posting and delivery of P3 student level data to the CDE	December and beyond, following the administration	Currently-offered post P3 data correction capability to be applied to subsequent student data file submission to the CDE

12. H Other Reporting Issues

12. H1. Longitudinally Comparable Scores

If requested by SBE staff, ETS will conduct a research study to investigate the development of a "Distance Between Predicted and Proficient" score that could be used to compare matched group performance from one year to another. This study will develop regressions for ELA and Math using 2004-05 state data, describe their statistical qualities and limitations, and provide some examples of their possible use in district level program evaluations. Matching of 2004 data with 2005 data will be done for students with ID numbers (either SSID or local identifiers). A report of the study will be provided to the CDE and the testing liaisons on the representativeness of the available matched data sample.

Each month during the study, a progress report will be submitted to SBE staff and liaisons and CDE. If requested, ETS will brief SBE staff and liaisons and CDE staff on the progress of this study prior to the delivery of the draft report. ETS will present, to the CDE and SBE, recommendations from the completed analyses (report). ETS understands that the contract may be amended at a later date to fund recommended activities and related charges.

This study will be conducted in two phases; Phase I will use 2004 and 2005 matched data. Phase II will be a replication of Phase I using available 2005–06 data. Phase II would only be

conducted if Phase I is successful as determined by the liaisons and the CDE. Additional costs to complete Phase II activities are contingent upon approval by CDE and SBE of a contract amendment and upon sufficient funds being made available by the Legislature in future fiscal years.

The proposed timelines and deliverables for these phases are described below.

Table 33. Research Study Deliverables

Deliverable	Due Date
Phase I (2004–5 Data)	
1) Approval of research study plan	1/6/06
2) Identify other states using regressions as growth measures; find out "lessons learned"	3/17/06
3) Examine existing data; develop appropriate rules for matching cases	1/20/06
4) Conduct regressions; determine technical details (e.g., linear vs non-linear)	2/10/06
5) Apply regressions to state level data and major demographic groups to identify anomalies or concerning results	2/24/06
6) Conduct several district level analyses as demonstrations (assuming sufficient data are available for one or more districts)	3/24/06
7) Provide a report for SBE testing liaisons and CDE, including recommend appropriate uses and safeguards for the growth measure	5/1/06
Phase II (2005-6 Data)	
1) Update plan and obtain approval to proceed	8/15/06
2) Replicate most important analyses with 2005-6 Data	10/20/06
3) Provide a report for SBE testing liaisons and CDE	11/3/06

12. H2 Shortening the Testing Window.

If the SBE determines through a change to Title 5 regulations to change the STAR testing window, ETS will plan with the SBE and CDE for the eventuality that the STAR testing window will shift to 90 percent of completion of the academic year and five days (rather than the current 10 days) on either side of that date. The ETS budget will specify cost implications to this proposed change in testing window assuming 1) that the cut-off date for receipt of answer documents for inclusion in the August 8 data release remains at July 1; 2) there is an realization that the actual number of districts and students included in the August 8 data release will vary from year to year depending on district adherence to the July 1 cut-off date; and 3) pilot district data review is expedited by using equating data from the "early form" used in the testing cycle.

13. Component Task 13: Reporting Test Results to CDE (CST, CAPA, CMA, STS, NRT)

The following annual reports and files are to be delivered to CDE by August 7, 2007 and to SBE if requested. The column labeled "State" in the following table shows the reports and research files to be delivered.

Table 34. Delivery of Summary Reports to CDE

		Distribution			
Report	**No. Copies**	**School**	**District**	**County**	**State**
County Group Summary CAT/6 Report	2			1	1#
County Student Master List Summary Report	2			1	1#
County Student Master List Summary EOC Report	2			1	1#
County Sub-Group Summary Report	2			1	1#
State Group Summary CAT/6 Report	1				1#
State Master List CST Summary Report	1				1#
State Student Master List Summary EOC Report	1				1#
State Subgroup Summary Report	1				1#
State Student Data File	1				1@
Internet Report	1				1@
Internet Research Files***	1				1@

@ Electronic copies preferred
# Paper and electronic reports
*** Multiple formats

State summaries are generated by grade and include all students in the State. The state also receives summary reports for each of the 58 counties and the California Youth Authority Schools.

13. A. Student Privacy

ETS will deliver the Internet site in accordance with these requirements.

- Data where a student's identity could be ascertained is suppressed and is represented by asterisks
- Reporting of all CST performance levels and provides a combined proficiency level which totals the sum of the Advanced and Proficient performance levels
- Allowance for the selective inclusion of either all five performance levels or the combined proficiency level on Web pages

13. B. Internet Site

13. B1. Requirements for Internet Site

ETS will work with the CDE to comply with the CDE's Internet standards. CDE will continue to host the STAR Internet site.

The site will be able to support 100 simultaneous users without significant design-related performance problems. To handle peak-demand or other times when the site is performing sub-optimally, static versions of all the Web pages are also supported. ETS will monitor actual Web site performance and work with the CDE to assure that the site meets performance specifications.

The design of the Web site will be data driven, so the user can very efficiently select the desired parameters to see the desired reporting of results. The database will use MS-SQL Server technology. While there are many combinations of summary reports that will be accessible, the summary data will be pre-calculated. While this may limit the dynamic nature of the site, it will prevent inappropriate summaries and interpretation of results by users.

The software application behind the Web site will allow the site administrator to load new iterations of data into the database and to generate new research files based on the refreshed data. As the data gets refreshed, notes added by the CDE from the previous iteration will be preserved.

The site will support two years of data: the current year and previous year. Summaries by counties, districts, schools, and the State will be provided. The site will support CST, NRT and CAPA as well as STS and CMA as those tests become operational.

13. B2. Delivery Aggregate Summary Data Files that are Synchronous with the Delivery of the Student Data Files

The Web site will provide for aggregate summary data files that are synchronous with the delivery of the student data files. These aggregate summary data files include aggregations by schools, districts, counties and the State. Independent charters are represented as separate districts within a county. The summaries will also be by individual assessment and by grade within each assessment. Courses that are not grade-specific will also be aggregated as an end-of-course summary. The summaries will include statistical data for the various assessments reflecting performance levels, quarters, or CAPA levels. These data will include the number of test takers, the average scaled score, and derived scores as appropriate. New STS and CMA results will be included as they become operationally available.

13. B3. Requirements for the Aggregate Summary Data

ETS will deliver report pages and research files that include aggregate summary data. The summary data and the Internet site will support the new requirement for CST cluster reporting. This cluster reporting will include such information as average percent-correct and mean-scaled score reported by grade (or by course for non-grade specific courses). ETS will work with the CDE to define this new requirement more precisely in order to optimize value to the STAR Program constituencies.

13. C. STAR Summary Data

The Internet reporting application will support the following demographic subgroups:

Table 35. Supported Demographic Subgroups — Internet Reporting

Demographic Category	Subgroups
All Students	
Gender	• Male • Female
English Learner 12-Month Status	• Mandated testing- students identified as "less than 12 months" • Optional testing- students identified as "12 months or more"
Special Education Services	• Students with Disabilities • Students with No Reported Disabilities
Economic Status	• Economically Disadvantaged Students • Non-Economically Disadvantaged Students
Special Program Participation	• Receiving Title 1 Services • Migrant Education • Indian Education • Gifted and Talented • EL in ELD • EL in ELD and SDAIE • EL in ELD and SDAIE with Primary Language Support • EL in ELD and Academic Subjects through Primary Language • Other EL Instructional Services • None
Ethnicity	• African American • American Indian or Alaskan Native • Asian • Chinese • Japanese • Korean • Vietnamese • Laotian

Demographic Category	Subgroups
	• Asian Indian • Other Asian • Filipino • Hispanic or Latino • Pacific Islander • Native Hawaiian • Guamanian • Samoan • Tahitian • Other Pacific Islander • White (not Hispanic) • Declined to state
Parent Education	• Not a High School Graduate • High School Graduate • Some College (Includes AA degree) • College Graduate • Graduate School/Post Graduate • Declined to State

13. D. Research Files

The Internet reporting application supports the following Research File requirements:

- State-level research file that contains all county, district and school results for all demographic subgroups
- State-level research file that contains all county, district and school results for the “all students” demographic subgroup
- State-level only research file that contains results for all demographic subgroups
- Limited research files that contain all data for selected counties, districts and schools
- A research file containing all CAPA data
- When available, research files containing all STS data
- When available, research files containing all CMA data
- A research file containing all reporting cluster results data
- Suppression of results where the reported group totals 10 or fewer students or where the number of student reports in any individual cell may allow identification of an individual student

- Compressed (zipped) research files formatted as fixed-length ASCII and comma-delimited (including column names) files. Provide an Access 2000 (or a more recent version of Access) database shell that can be used to import comma-delimited research files along with all instructions for use of the database shell. Provide a load utility that will facilitate the easy importation of comma-delimited research files into the database shell.
- Starting with the 2007 STAR administration, ETS will deliver Web-based data to the CDE in the XML format. The specifications for this data file will be part of the annual report specifications process.

13. E. Administrative Functionality

ETS will incorporate extensive administrative functionality into the Internet design to include:

- **Notes.** Allow for the inclusion of "notes" that may be dynamically added to any selected report page. For example, notes may be added to one or all schools in a district and to one or all of the subgroups. Notes must be capable of being retained when report data is updated.
- **Embargo Reports.** Allow for the selected exclusion of Internet report pages. For example, all cluster reports may be excluded, or the extended proficiency CST report page may be embargoed for subgroup reports at the school level while the combined proficiency report (combined total of proficient and above students) is accessible. Also, all state reports are embargoed until the site is opened to the public.
- **Research File Generation.** Allow for the generation of new research files when new aggregate data is loaded to the site. Which files are generated and the sequence of that generation must be part of the research file generation function.

13. F. CDE Web Delivery Requirements

Starting with the 2007 STAR administration, ETS will deliver Web-based data to the CDE in the XML format. The specifications for this data file will be part of the annual report specifications process.

The key to successful deliveries of the Internet reporting application and data files is to plan for preliminary iterations. This strategy allows CDE data management staff to be involved in early review of the site and the data. By delivering early, issues are identified and remedied earlier, before the critical public deadlines.

ETS proposes the following timeline for site development and data deliverables:

Table 36. Timeline 4: Proposed Timetable for Site Development and Data Deliverables

Month	Deliverable
January	ETS works collaboratively with the CDE to document business requirement changes and other changes for the Internet reporting application.
February	ETS updates specifications.
March	ETS makes modifications to the site, including initial content/text changes provided by the CDE.
April	ETS deploys the site in the ETS user acceptance environment for the CDE's first review of the site. ETS loads the prior years' data into the site (updating the format and content of the data files as necessary).
May	ETS deploys a second release of the software, based on CDE feedback as well as internal testing results.

Month	Deliverable
June	ETS deploys a third release of the software to the user acceptance environment as needed and posts early production test results (referred to as the V1 deliverable in the current contract). ETS also delivers aggregate and corresponding student data files to the CDE. ETS also delivers the application to the CDE for installation to meet the required July 1 deadline.
July	ETS deploys additional releases in the ETS user acceptance environment as necessary (before the July 13 deadline) to correct defects and make final text/content changes from the CDE. A significant amount of production data is now loaded into the site, representing approximately half of the testing population (two million plus test takers, i.e. V2 data release in the current contract). ETS delivers the V2 aggregate and student data files to the CDE.
August	ETS delivers data files and software to the CDE by the August 4 deadline. The aggregate and student data files delivered in early August represent the P1 deliverable in the current contract. ETS will also load the P1 data to the user acceptance site, to stay synchronized with the CDE. P1 data includes all districts that completed testing and returned results by a certain date to be determined by the CDE (June 30 in the current contract).
September	ETS delivers P2 aggregate and student data that now include all districts. ETS will also load P2 data to the user acceptance site to stay synchronized.
December	ETS delivers P3 aggregate and student data that now include all districts. ETS will also load P2 data to the user acceptance site to stay synchronized.

Provision of Web-based data in XML format.

Starting with the 2007 STAR administration, ETS will deliver Web-based data to the CDE in the XML format. The specifications for this data file will be part of the annual report specifications process.

ETS will make every effort to assure that all Web development will comply with the IT Compliance clause of the STAR contract.

13. G. Secure File Transfer

Due to the confidential nature of test results, ETS uses File Transfer Protocol (FTP) and encryption for all student data files. In addition, ETS uses Zip technology to reduce the disk space requirements on all files. This method applies to all data file transfers.

13. H. Student-level Files

ETS will deliver student level data files on dates set by the CDE.

ETS will deliver one or more snapshots of Pre-ID data to the CDE on a schedule to be specified by the CDE.

ETS will deliver student data files and corresponding aggregate files on the delivery schedule in Section 13.F.

ETS will deliver student data files in three formats: a compressed layout with demographic information only, a layout with item response data and demographic information, and a file that contains all student data available.

14. Component Task 14: Technical Report, Other Reports, Analyses, and Data Collection (CST, CAPA, CMA, STS)

14. A. Technical Report

ETS will develop, maintain, and provide to the CDE and the SBE all documentation needed to assure the technical quality and continuity of the CSTs, CAPA, CMA, and STS including, but not limited to Technical Reports. The Technical Report will document all aspects of developing the CSTs, CAPA, CMA, and STS.

All narrative reports submitted by ETS will include an Executive Summary, the full text, and appendices containing all relevant data tables. In addition, the Executive Summary will be written to stand alone as a document suitable for public distribution. All final narrative reports and all electronic deliverables will be provided in MS-Word, PDF, and HTML format for distribution and possible posting on the ETS STAR Web site. All tables and technical appendices will also be submitted in MS-Excel spreadsheet versions.

Annual Technical Reports will be assembled for each testing program (CSTs, CAPA, CMA, and STS) from the equating and P2 file data and will be delivered to the CDE no later than November 1 of the testing year and at the termination of the contract. The CDE will have 20 working days to review the Technical Report. Five bound paper copies of each Technical Report will be submitted annually to the CDE, as well as one copy on CD-ROM in PDF format.

The Technical Report will be supplied as a MS-Word document, and will be organized and clearly labeled to facilitate cross-reference to the Standards for Educational and Psychological Testing (1999). Sections of the Technical Report will be written by specialists in their respective areas of concentration.

To support the CDE in providing evidence of meeting the requirements of *NCLB*, ETS will provide Technical Reports that clearly identify critical elements for *NCLB* Peer Review addressed within the scope of the reports. In addition, ETS will work hand-in-hand with the CDE to develop evidence obtained over more extended periods; for example, the development of validity evidence to support each testing program.

Tables and figures will be included in the Technical Reports as needed to summarize and clarify analysis results and development procedures. Further data analyses for the purpose of assuring the validity of test scores, federal peer review, programmatic review, program evaluation or any additional inquiries regarding the operation of the STAR, CMA, STS and CAPA assessment programs will be readily provided by ETS. ETS will also discuss additional research studies to support the STAR Program. Results of any additional analyses conducted at the request of the CDE or the SBE will be included as well.

14. B. Student Information Report for Apportionments and API

This Student Information Report data is readily found in the student data file. ETS will deliver this summary report to the CDE with each iterative deliverable of aggregate and student data for each administration year. ETS will coordinate with the CDE on the format of this new deliverable.

14. C. Instructional Material Survey

ETS will annually create an electronic data file for the Instructional Materials Survey that is included on the School Grade Identification (SGID) sheet. ETS is to forward the file on CD-ROM to CDE two weeks after the delivery of the P2 file each year using the following file layout:

Table 37. STAR Instructional Materials Survey Record Layout

Field First Position	Field Last Position	Field Length	Field Description	Field Acceptable Values
1	2	2	County Code	
3	7	5	District Code	
8	14	7	School Code	
15	23	9	Order Number	
24	63	40	District name	
64	103	40	School Name	
104	105	2	Grade	Alpha (02 – 11)
106	135	30	Reading Language Arts	Numeric (1) or Blank (b) 1 = Yes b = No Response (Bubble 1 is first position of array, Bubble 2 is second position of array, etc.)
136	165	30	Mathematics	Numeric (1) or Blank (b) 1 = Yes b = No Response (Bubble 1 is first position of array, Bubble 2 is second position of array, etc.)
166	195	30	History	Numeric (1) or Blank (b) 1 = Yes b = No Response (Bubble 1 is first position of array, Bubble 2 is second position of array, etc.)
196	225	30	Science	Numeric (1) or Blank (b) 1 = Yes b = No Response (Bubble 1 is first position of array, Bubble 2 is second position of array, etc.)